A STUDY OF THE FACTORS CONTRIBUTING TO INTER-TOWN DIFFERENCES IN INDUSTRIAL STRUCTURE

N. R. GILLHESPY

Ph.D. UNIVERSITY OF EDINBURGH JUNE 1969

ACKNOWLEDGEMENTS

I wish to acknowledge the general supervisory help extended to me by Professor J. N. Wolfe.

In addition I wish to express my appreciation to Professor P. Vandome for his continuous advice and criticism of my empirical section. I would also like to convey my appreciation to W. D. C. Wright for some most timely comments, to A. J. Parsley for his advice on the use of computer programmes, and to Miss J. L. Milne for preparing the diagrams.

CONTENTS

APPENDICES

SUMMARY

Past regional studies have been unable to predict those changes in industrial structure consequent upon developments in the regional economy and the broad aim of this thesis is to secure an improvement upon this situation. This is approached in two stages. First a review of those theories relevant to forecasting industrial structure is undertaken and, secondly, in the light of the conclusions so drawn, a theoretical model is derived and then tested.

It is a characteristic of the first step that there is no definitive work on the subject and it is necessary therefore to cast the analytical net over a wide area in order to encompass all those theories which may lead to useful hypotheses. Attention is initially attracted to location theory, for this represents the only attempt at deciding general laws about the location of industry, and, therefore, the industrial structure, of any one region. This approach, however, has serious deficiencies at a practical level and this has led subsequent writers to increase its usefulness through the development of less ambitious techniques. These are mainly based upon the notion of the presence of an hierarchy in regional units, so that the function (structure) of each unit will be decided by its position in the hierarchy. From these developments two hypotheses are taken, one based upon the regional unit's population and one upon its geographical location.

An alternative set of theories based upon macro-economics is found to be available to the analyst. However, at this stage, it is necessary to introduce a further characteristic of this study, viz. the

lack of sufficient data. This has the effect here of reducing the range of theories available to the Sector and Regional Multiplier theories. These then lead to two further testable hypotheses, one based upon the region's external economic connections and another upon its own per capita income.

These four hypotheses can then be used to provide the basis for a theoretical model aiming at specifying the factors likely to influence industrial structure. First, however, 'industrial structure' must be defined. An approach based upon individual industrial orders was rejected on the grounds that the techniques are not sufficiently sophisticated to allow for such a level of disaggregation, and the broad primary, secondary and tertiary sectors are adopted as the principal components of industrial structure. Due to the smallness of the primary sector, the main division must lie between the secondary and tertiary sectors, and here the model has been cast in terms of predicting the size of the tertiary sector.

Adopting this definition of industrial structure, the four hypotheses are used to derive a general model which a priori could be applied to a wide variety of situations. However, before further progress can be made, the data restrictions must be explicitly introduced and the general model accordingly modified. The latter step is based upon the key assumption to take the town as the basic unit of analysis. A further difficulty with the data is the lack of any measure of per capita income at the town level. Several proxy variables are suggested, but only two, non-domestic rateable value and an index of earnings, are found to be acceptable.

The model so defined is now tested. The sample chosen consisted of all the towns in the British Isles with populations in excess of 50,000 after deleting those within a 50mile radius of London or within officially defined conurbations. This resulted in a sample of 69 towns and these were analysed through a cross-section regression analysis. Only three hypotheses are found to be significant. One of these, the geographical location variable, presents no problem, but the other two carry, from a theoretical point of view, the 'wrong' sign. The first of these is the above mentioned index of earnings, whilst the second stems from the regional multiplier analysis.

A further look at the model is taken with a view to (a) explaining the negative coefficients, and (b) improving upon the variance explained by the theoretical model. The latter step involved investigating new (pragmatic) hypotheses suggested by the behaviour of the data, together with considering how the remaining theoretical variables could be refined. The new hypotheses are then tested and the significant ones integrated with the three theoretical variables. The final model explained 82% of the variance and contained the index of earnings, the derivative of the regional multiplier theory, the degree of commuting by the labour force from outwith the town and an 'historical' element as the independent variables. The last influence was represented by the size of the tertiary sector 1931 and an index of hotel accommodation.

Chapter One

INTRODUCTION

The broad field of enquiry with which this thesis is concerned was first suggested by a practical problem met in regional economic development. The Government in its plan for the Scottish Economy (49) proposed to increase the population of the Central Borders from 75,000 to 100,000 over the period 1965-1980. In the face of such a population increase one of the first issues which a regional economist must consider is the likely impact that this will have upon the region's industrial structure. For instance will the whole of the population increase find employment in an expanded manufacturing sector or will some pass into the service industries? If the latter is the case then what can be said about the relative growth rates of the two sectors? Moreover, within the manufacturing sector, will the increase in employment take place solely within the existing industrial base or will new industry be attracted into the region and if so what type will it be?

These questions are of great importance to the regional analyst planning for such a population growth whether he be an economist, town planner or architect. Before the economist can assess for 1980 what will be the region's income, the composition of its demand for labour, its relationship with other regions etc. he needs to know something about what form the industrial structure will take by the end of the expansion period. At a more sophisticated level he will need the same information in order to predict the future growth of the region, its stability with regard to national cyclical fluctuations and its inter-industry relationships. Furthermore the future industrial structure is no less important to the town planner or architect for without it he cannot state what size the town centres will have to be, how much land should be reserved for industry, what water, electricity and other ancillary services should be provided etc.

Although the problem thus posed stemmed initially from the Government's plan in connection with the Central Borders, it is nevertheless one of the first issues that an economist analysing the development of any region will have to answer. That being the case it is natural that the next question to be asked should be how can future industrial structures be forecast or, to put the point another way, how have past development projects dealt with this problem? It is somewhat surprising to find, therefore, that economic theory has been notably reticent about the whole problem of forecasting industrial structures, a fact which naturally enough has been carried over into the various development reports. For instance, a fairly typical conclusion on the simpler issue of the division between the secondary and tertiary sectors is to be found in the Lothians Regional Survey and Plan (48).

> 'The changes which have occurred in the Survey Area since 1961 have probably resulted in approximately equal numbers now being employed in manufacturing and in service industries. Thus the range of variation in employment provided for the planned increase in population will be from 60% in manufacturing and 40% in service industries to 50% in each.' (p.61)

Of the various specific development studies Dunning (13) perhaps provides the most satisfactory analysis for he attempts to reduce the possible variation by classifying industries into a) export industries, b) service trades linked to the export industries and c) trades catering for local consumption. Nevertheless the analysis still yields a result no more accurate than that the size of the secondary sector is likely to be determined by 'a yardstick of 45-50 per cent' [(13) p.94].

In order to find any theoretical attempts at forecasting industrial structure one must go beyond the purely empirical case-studies approach and into the formal body of economic theory itself. But it would appear to be a characteristic of this problem that even then no one single body or theory can be identified, but rather that the various theoretical approaches are to be found scattered over a wide field. It is therefore necessary to categorise the theoretical writings on this subject and to this end, at the risk of over-simplification, a six-fold classification may be effected.

The first group consists of the relative income potential models which have been used to project the growth of a region's economy. Isard (27) states that two of the reasons why a region may grow are a) it is part of a national economy which grows or b) relative to other regions its access to its markets improve. Isard bases his analysis upon 'market-access sensitive' industries and analyses the growth in population in terms of two effects upon these industries. The 'proportional effect' measures the influence that growth in the economy has upon the region under review, whilst the 'differential effect' indicates the influence that changes in access to its market have upon the region.

The chief disadvantage of this method lies in its concern with forecasting population (though suitably amended the model could also be interpreted in terms of employment) and its consequent lack of interest in any detailed analysis of industrial structure. In other words it can be used to forecast a region's overall industrial activity and as such it is a useful tool, but it is unable to predict the size of the various elements of this activity.

The second category, rather than studying the regional economy in general, goes to the other extreme and considers various industries which may form a part of that unit. One classic example of this approach is to be found in industrial complex analysis as propounded by Isard and Schooler (28) in which, on the basis of output matrices, meaningful industrial complexes can be identified. Then, on the basis of cost comparisons, either the optimal location for plants or alternatively the best type of plant for a given location can be determined. However in practice it is clear that the number of plants or the number of locations for which this analysis could be applied must be very restricted restructed. As a result, information of this type would be valuable background knowledge for a region, but it could never be assembled in sufficient detail to identify by itself the future industrial structure of a region.

A less ambitious, but nonetheless more useful, approach lies in the attempt to identify the size of a region's or town's retail sector. In this method Reynolds (45) determined a town's theoretical shopping population, or its drawing power, either through the gravity formula or rateable values, and then plotted this population against the size of the retail sector. A step-wise function was found indicating that the retail sector does not increase appreciably until certain theoretical shopping populations are met. A similar result has been reported by Clark (9) for manufacturing industries in New Zealand. Nevertheless, although this method can be applied to any region, it still fails to provide more than essential background information to an analysis of industrial structures, for the simple reason that retail trade is but a small part of the total industrial structure.

Into the third category fall the various attempts to postulate econometric models to explain past changes in industrial structure. These models, in an attempt to keep the relationship consistent with the data available, generally take as a central tenet the assumption that manufacturing employment is a function of population. On the basis of this relationship the models are used to explain the distribution of industrial activity either geographically or economically. As an example of the former Niedercorn and Kain (38) consider the distribution of manufacturing employment between the central city and the metropolitan ring, whilst for the latter a typical example would be Czamanaski (11) in which industrial activity is divided into geographically orientated, urban orientated and complementary industries. Obviously both these approaches are of use in their respective fields and the latter may have considerable value in gaining a rough guide of a region's industrial structure. Yet, because these approaches rely upon the simple relationship between population and manufacturing employment, as a reliable test for forecasting the future industrial structure in detail they progress little further than simple extrapolation.

One approach which avoids this difficulty is that of the two-sector model as originally developed by Uzawa (61). The use of this model has theoretical appeal not only because it recognises two entirely separate sectors of the economy, but also because it is involved with the dynamic process of growth, a point which is of clear relevance to the forecasting of industrial structures consequent upon an expansion in the regional economy. Unfortunately, as with most econometric models as yet devised, in order to progress theoretically a sacrifice in realism is called for which, when the model is re-applied to a specific practical problem, is generally unacceptable. For instance in Uzawa's models the rigid dichotomy between the capital and the consumption sectors is never found in practice. Moreover the model is concerned with determining the conditions necessary for a steady-state and to see whether the equilibrium path will tend towards this. However since most regional economic development itself involves the creation of a disequilibrium position, this is but further evidence of the lack of realism of this approach.

The fourth category, that of input-output techniques, offers far more scope for practical progress. Basically regional input-output analysis consists of a matrix of input-output coefficients relating not only to inter-industry but also to inter-regional flows as well. On the basis of an assumed rate of growth in either final demand or the national economy the relative expansion paths of each industry in each region can be calculated and, therefore, the change in the region's industrial structure forecasted. Clearly in its idealised form the input-output table's data requirements are too sophisticated and Meyer (35) has illustrated the principal ways in which the situation may be eased. Attempts have been made to aggregate the flows either by subsuming all inter-regional flows into one import/export sector or, conversely, aggregating all industries and concentrating upon inter-regional relationships. On the basis of these simplifications a considerable number of successful attempts at applying input-output techniques to specific development problems have been made [see for instance the development of an inter-industry model of Utah by Moore and Paterson (36)].

It would be idle to deny the considerable progress made in forecasting industrial structures by the adoption of input-output techniques, but there are nevertheless certain drawbacks to the system which make it not unjustifiable to search for alternative methods. First, regional input-output tables must take the national economy as the starting-point and they work downwards towards the regional economy. This may be a legitimate exercise if the region is large, for what is true at the national level is then likely to be true also at the regional. But if instead of considering Scotland or Wales only a part of the region is considered, say Fife or North Wales, the method becomes less satisfactory for too many industries are not represented and too many national coefficients become susceptible to variation as a result of strong, but purely local, influences. In other words the analysis is too inflexible to be used in practice in regions which are other than a substantial sector of the national economy. Secondly, the technique is purely descriptive, it depends upon past relationships and it contains no behavioural element. In the short term this may not prove a serious disadvantage but since in the long term the success of regional economic development depends upon changing local consumption and production patterns, the relevance of its conclusions must be seriously curtailed. In other words so long as the regional change is associated with a similar overall change in the national economy, input-output analysis is a distinctly viable technique, but when the change emanates from Government or other outside manipulations or from purely local influences the analysis must be treated with extreme caution.

To summarise the position so far, the relative income potential, industrial complex and retail sector models all form valuable background knowledge for an understanding of a region, but they fail to consider the central problem of industrial structure itself. The various mathematical formulations found in the third category are generally too simplified and too unrealistic for any direct application to the problem, whilst input-output analysis is too inflexible and lacks any behavioural content to render it appropriate to any region which is not of a considerable size in relation to the national economy. Ideally a study of industrial structure would require a method which is applicable to a substantial sector of the region's economy and not just to one or two industries with convenient characteristics. In addition the method should contain a certain behavioural element so that it does not depend entirely upon past relationships and/or coefficients and could consequently be used in situations where typically these are altered. Finally the method should be reasonably flexible in order to enable it to be applied to a variety of types and sizes of regions. Clearly these conditions are idealistic but they nevertheless set standards against which the alternative theoretical formulations can be assessed. As already noted the methods so far reviewed perform somewhat inadequately when set in this context and it is the contention here that the remaining two categories to be discussed offer more scope.

Into the first of these two categories fall location theory and its descendants. Location theory is the only general theory which explicitly takes as its subject matter the formulations of general laws dictating the course of industrial location. As such it provides an analysis which can be applied to any or all parts of the regional economy and since it aims at general laws its conclusions are correspondingly capable of application to a variety of divergent situations and different types of regions. Similarly its descendants, whilst adopting rather different aims and methods, still retain most of the essential features of location theory.

The second category consists of those purely macro-economic theories which can be applied directly to the problem of industrial structure. There are two such theories, both of which concentrate upon the demand generated within a regional economy. The first of these is the sector theory in which the size of the tertiary sector is related to the region's income. This method is less satisfactory than location theory or input-output theory in the sense that it analyses only broad sectors of the economy, but nevertheless the division between secondary and tertiary sectors is still of considerable importance to the regional economist. But the method has considerable behavioural content and a priori is capable of application to a variety of situations. An alternative to this demand approach is to be found in the regional multiplier and economic base theories in which changes in the exogenous sector are traced into the endogenous sector. It is true that to a certain extent the method lacks behavioural content inasmuch as the value of the multiplier must be assumed constant, but these theories are capable of widespread application and represent the most flexible tool of analysis open to the regional economist.

AIM OF THESIS

It is clear from the above discussion that in the case-study approach the important question of a region's future industrial structure has received somewhat inadequate treatment and a general desire to be able to state something more precise about industrial structures in the future forms the main concern of this thesis. It has already been noted that it is a characteristic of this subject that the various writings which could offer possible tools of analysis are cast over a wide field. This, together with the absence of any definitive work on the subject, suggests that the method finally selected should be based upon a synthesis of the most promising aspects of past writings.

From an initial and cursory review of the deficiencies of past formulations it is clear that one of the most important of these is to be found in their operational inflexibility, and that this basically stemmed from their reliance upon parameters developed originally for the national economy. In order to overcome this, any future technique for predicting future industrial structures would have to be 'regional' in the strictest sense of the word, i.e. it would have to be a technique specifically evolved in a regional context which was independent of national parameters. A second important weakness of some of these approaches resulted from their need to rely upon past relationships. This suggests that any future method should contain a greater behavioural or motivational element so that it is based not upon relationships which are true only at one point of time, but rather upon general patterns or tendencies (law is perhaps too strong a word) which can reasonably be supposed to be constant over time. Although to overcome the deficiencies entirely is clearly an impossible task it was nonetheless suggested in the last section that location theory and its descendants, together with the sector and/or regional multiplier theories, offered the greatest scope in this respect. It is therefore proposed that these theories should be taken as the theoretical starting-point for this thesis. Then, from a study of these, a model may be derived representing the most viable ma. of their propositions, to forecast industrial structure.

SCHEMA

This thesis is divided into two sections, the theoretical and the empirical. In the former, location theory (chapter 2), its descendants (chapter 3) and the borrowings from macro-economic theory (chapter 4) are critically reviewed. From the discussions in chapters 2-4 a general model for forecasting industrial structures is derived, from which a particular model is developed in order to meet the data requirements of the chosen field of enquiry (chapter 5). In the empirical section this model is then tested. In chapter 6 the implications for the empirical research of the data used are discussed and in chapter 7 the actual research findings are set out. Finally chapters 8 and 9 consider the weaknesses of the model developed in chapter 5 and attempt, through suitable modifications, to overcome these.

THEORETICAL SECTION

Chapter Two

LOCATION THEORY

Although location theory itself cannot be applied directly to the problem of forecasting industrial structures, it is nevertheless necessary to open this theoretical section with a study of this theory. The reason for this lies partly in the fact that the theory represents the only attempt to explain explicitly the location of economic activity. *Prima facie* it is therefore of great relevance to this thesis and to exclude its use certainly requires some justification. But the reason for considering location theory goes further than this, for many of the subsequent theories to be incorporated in the final model find their origins in the workings and deficiencies of location theory and in order to fully appreciate these theories it is necessary to view them in the context of their relationship with that theory.

It is therefore the purpose of this chapter to review the contributions made by location economists, to develop those ideas which form an essential background to the subsequent chapters and, finally, to show why location theory itself cannot be applied directly to the problem of forecasting industrial structures. On this basis the chapter can be divided into two parts, the first outlining the general principles of location theory, the second being devoted to an assessment of its applicability to the problem at hand.

REVIEW OF LOCATION THEORY

The aim of location theory is to study what factors influence the location of economic activity and changes in that location. In so doing location economists have tended to force a dichotomy between their subject matter and that of the rest of economic theory, but as will be shown from a review of Isard's work (26) this division is largely, though not entirely, false. This notion of the separate nature of location theory was largely due to Weber (63) who gave the theory a narrow substance through his insistence that labour and transport costs were the only two costs of location and also by his manner of solving location problems through mechanical models. As a fair generalisation the former concept is still present today, whilst the latter has undergone considerable sophistication.

Weber postulates that labour and transport costs are the only two general factors that vary with, and thereby influence, the location of economic activity. There were others, principally the secondary factors consisting of agglomerating and deglomerating forces (see below) but these themselves are a result of the primary factors. Given the presence of a) an even plain with equal transport costs and b) an equal cost of inputs at their source, it follows that the location of economic activity will be determined by the interplay between the ratio of the weight of localised material to that of the product (the 'material' coefficient) and the ratio of the cost of labour per ton to the total weight of the product (the 'labour' coefficient). In the absence of any mathematical formulation, Weber had to rely upon mechanical devices based upon weights and strings to determine the precise location of any individual industry.

The other main contribution made by Weber was his development of the notion of an hierarchy in economic activity. Weber considered the forces which were brought into play when an undeveloped country was occupied. He indicated the formation of five separate strata viz. 1) agriculture, 2) primary industry, 3) secondary industry, 4) central organisation and 5) central dependant. In his analysis of locational structures the five strata are all inter-related with each one determining the loci of consumption of the one above it.

The first major advance to be made upon Weber's position came from Predöhl (43) who considered location theory as a special case of the theory of interdependent prices and quantities as expounded by Walrus, Pareto and Cassels and thereby treated the location problem as merely one of price. This approach had the advantage of, first, infusing more economic flavour into hitherto purely technical considerations and, secondly, of replacing the crude mechanical methodology with a mathematical formulation. Thus, as a result of Predöhl, location theory became equipped with a framework in which economic variables could be introduced and allowed to alter. But although Predöhl based his theory upon the Walrus-Pareto-Cassels theory he nevertheless failed to develop the analysis to its logical conclusions viz. reducing the theory to a set of equations and it was left for the next major contributor, Lösch (32), to do this.

However, as far as this thesis is concerned, the major interest of Lösch's work lies not in the novelty of his methodology but rather in the originality of his conclusions; namely that even in the absence of any special historical or geographical explanations a logical pattern of location would develop and that this pattern would manifest itself in the form of an hierarchy. As will be seen from the next chapter these conclusions represent the essential theoretical foundations to central place theory (which will be used extensively in the specification of the final model). These conclusions can be arrived at in two stages. First Lösch deduces that economic activity must always be spatially differentiated as a result of the combined interaction of economies of scale, transport costs and agriculture's need for space. With economies of scale, but without transport costs, all production would be concentrated at one or two randomly located optimum sized plants, whilst in the reverse situation production would be totally fragmented. With both, but without the ingredient of space, the result would be a compromise between mass production and freight costs. But since space itself is an input to agriculture, this fact alone is sufficient to force non-agricultural production into smaller sites. Economic activity must therefore be spatially differentiated.

The second stage is to identify what pattern this spatial differentiation will take, but before this can be achieved it is necessary to define the 'ideal market'. This Lösch derives from the postulates of a) an even distribution of raw materials and population over a homogeneous plain and b) the accessibility to everyone of all occupations and all methods of production. Each producer will then be faced with a falling demand curve consequent upon the need to increase prices as the distance over which he sells increases. Since the producer sells in all directions the total demand confronting him will be a cone (obtained by rotating the demand curve about the price axis). His market is the area over which he sells and this is represented by the base of the cone i.e. it will be a circle. However, under this arrangement there would be a space between the individual markets thus allowing one firm to expand and to reap excess profits. Consequently the final pattern of markets will be a regular hexagon since, out of all the forms that are capable of spatially covering the whole area, this deviates the least from the circular, thereby minimising transport costs.

This then is Lösch's definition of an ideal market, but of particular importance now is what the overall pattern will look like. Here Losch puts the point most succinctly and one can do no better than quote him at some length. He states (32) p.73

> 'The trading areas of the various products look like nets of such hexagons... We can throw these nets over our plain at random. In spite of the resulting disorder, every place on the plain would have access to every product. Several considerations ... suggest a more orderly fashion. In the first place, we lay our nets in such a way that all of them have one center of production in common. This point will enjoy all the advantages of a large local demand. Secondly, we turn the nets around this center so that we get six sectors where centers of production are frequent and six others where they are scarce. This arrangement does not deprive any place of its access to every product and at the same time provides for the best lines of transportation... More striking about our result than any particulars is the fact that we suddenly have crowds of economic areas on a plain which we deprived of all spatial inequalities at the outset. We first have the hexagonal market area surrounding every center of production or consumption. Second we have a net of such areas for every commodity. And third, we have a systematic arrangement of the nets of market areas of the various commodities.'

Thus Lösch is able to deduce that economic activity will not only be arranged in an hierarchy but that each sub-group will be spaced a definite distance apart not only from the other sub-groups but also from the ones above it. But as Stopler (55) has noted 'while the pattern of location itself and the relation to the market area is determinate, the location of the whole pattern in the plain must be indeterminate' (p. 632). In other words Lösch can deduce the position of each component in his system without resorting to historical and geographical explanations, but it would appear that such explanations are still required to determine the central point of the whole system.

Up to this stage in the development of location theory the contributions have been couched in terms peculiar to each author, which has resulted in the theory appearing to be far removed from the general body of economic thought. However Isard (26), the final writer to be considered here, manages to integrate much of the past developments with the rest of economic theory and he achieves this by restating the previous conclusions within the framework of the substitution principle. Thus, given the assumption of profit maximisation, producers will choose, *ceteris paribus*, that location which minimises total location costs, a point determined by an orthodox substitution process. As a result, Isard manages to illustrate how spatial considerations may be incorporated into traditional economic theory by the simple process of expanding that theory so as to take cognisance of those costs which are dependent upon location. The basic unity not only between location theory and traditional theory, but also within location theory itself, is now apparent, though Isard would appear willing to go even further than this in his claim that:

> 'modern general equilibrium theory is a special case of this (location) theory in which transport costs are taken as zero, and all inputs and outputs are viewed as perfectly mobile' (26) p.53.

Yet in some ways this is a somewhat over-stretched juxtaposition of the two theories and an oversimplification of their differences, for it presumes that all the contributions of location theory can be re-cast in the mould of the substitution principle. However, as far back as Weber it was realised that it was possible to effect a two-fold division in the location factors viz. primary and secondary factors. The principle can be applied easily to the first but with the secondary factors, principally the agglomerating and deglomerating forces, the principle quickly runs into difficulties.

Hoover (23) has distinguished three types of agglomeration economies, i.e. those economies which cause firms to locate together. These economies are, first, economies of scale. Secondly, there are the localisation economies which accrue to all firms in a particular location as a result of an increase in total output at that location (e.g. those economies associated with common pools of skilled labour, fuller use of specialised services etc.). The final category, urbanisation economies, consist of those economies which accrue to all firms in all industries at a single location consequent upon an increase in economic activity at that location. The deglomerating economies, which roughly equate with the three agglomerating economies above, are 1) diseconomies of scale, 2) the rise of both rents and the cost of urban services as land use intensity increases and 3) the rise in the cost of food and other services as a fuller use is made of the area's own resources. These opposing influences are usually netted out and the resultant force expressed as a function of agglomeration or deglomeration.

The first of these, economies of scale, can easily be incorporated into the substitution framework, since all this means is that entrepreneurs have now the extra alternative of substituting a decrease in production outlays with an increase in transport outlays. In other words the scale of output joins transport and labour costs as one of the fundamental basic decision variables. Although in principle entrepreneurs could substitute the second category, those of localisation economies, such a procedure would be valid only in the situation where the location was on a completely new site. But usually there is already in existence a physical structural framework and to relocate plants away from this would involve the opportunity costs of forced plant obsolescence. Further, once one production plant has already been established others are likely to locate there so as to take advantage of the localisation economies present. Hence the evolutionary framework now must become a location factor and it is obvious that this could never be susceptible to substitution analysis. Although the third category of urbanisation economies includes these localisation economies, they are even wider and therefore even less susceptible to substitution analysis, since they also embrace a) those economies which stem from a greater use of the capital infrastructure and b) the diseconomies engendered by a rise in the cost of living, wages, rents etc.

Consequently, although Isard manages to unify and make intelligible large tracts of location theory, it is still not possible to consider that theory as falling entirely within the ambit of traditional theory. Moreover since the agglomerating and deglomerating factors not only fail to fit into the substitution scheme but also fail to be susceptible to any other form of analysis, one aspect of location theory must not only be thought to be alien but also to be rather unsatisfactory.

AN ASSESSMENT OF THE APPLICABILITY OF LOCATION THEORY

From the above résumé of location theory it can be seen that the main aim of the theory has been to derive general laws on those forces which influence industrial location. As such the theory might be thought to be tailor-made to the problem of forecasting industrial structures, for clearly what such a problem requires is a study of those factors which would influence the location of industrial activity. Moreover its applicability is further enhanced by the theory's pre-occupation with a general solution. Yet any attempt to apply the theory directly to the problem at hand falls down on account of practical weaknesses concerning its content, methodology and scope. In the following it should be noted that these weaknesses are associated with the whole of location theory and that as a result it is not necessary to look to those criticisms concerning the individual theories in order to find grounds for rejecting this approach.

With regard to the theory's content, despite Lösch's formulation of his ideal market, location theory is principally concerned with costs. But as Greenhut (20) rightly points out:

> 'In the instances where buyers are dispersed the selection of a plant site involves more than just the minimisation of costs of sales to some given buying point. Demand, in fact, becomes an important variable which depends upon location and which actually may be more variable than cost from place to place'. (p.176)

The reason why demand is such an important variable lies in the fact that it causes, and at the same time reflects, varying uncertainty and profits and that, therefore, it influences and is influenced by the location of industry. A second criticism that can be made about the content of location theory is that it fails to develop satisfactorily a theory about those costs which are important. Certainly it would appear that its emphasis on transport costs is misplaced, for the Toothill Report (47) found, on the basis of a survey conducted amongst 95 firms in Scotland, that 87% of the firms had transport costs less than 3% of total costs. Further, out of 45 firms that had experience in both Scotland and the south only 26% found the extra transport costs more than 1%. Much controversy has centred around the importance of transport costs and perhaps the most reasoned position is that adopted by Clark (10) in which he points out that some industries are market or material-orientated (i.e. their location is determined by transport considerations), but that the remainder of the industries, comprising in fact the vast majority, are 'footloose'. It would surely be more important to determine what forces dictated the location of these industries. Location theory, on the other hand, is on firmer ground with its emphasis on labour and agglomerating factors. But in this respect it is a pity that agglomeration analysis, in the words of Isard (27), 'has little to say beyond the obvious: units are attracted to or repelled from cities according to a simple comparison of advantages and disadvantages'. (p.183)

The criticism about the methodology adopted by location theory rests on the fact that the essential assumption behind the substitution principle has been shown to be unrealistic. The whole solution proposed by location theory relies upon the fact that entrepreneurs will substitute one cost for another so as to arrive at the optimum location. To refute this assumption is to refute the use of the substitution principle. Yet, in a survey based on 200 firms that had re-located away from Birmingham, Loasby (31) found that

> 'the typical firm's first idea is to find something close at hand, and only if this proves impossible does it begin to look further afield; it then searches outwards and accepts the first site which seems satisfactory. It does not worry about considering all the possibilities, but wants a solution quickly and easily, and quickly and easily implemented' (p.36).

The suspicion that firms do not consciously evaluate the alternative sites open to them is corroborated by the findings of Trotman and Dickenson (59), based on a survey of 21 Scottish Industrial Estates, that few firms had studied the cost of production in other locations before choosing their own Scottish site.

The final criticism to be made about location theory concerns its scope. The originality of the theory lies in the breadth of its scope. Yet, from the point of view of a practical application, this robs the theory of much of its use. Even if it was possible to determine what general factors govern the location of industries, it would be still virtually impossible to use the conclusions as much more than guidelines. This is so not only because each region has its own important, but unquantifiable, characteristics but also because each town has a different role to play within its region. Thus what may be true for all regions would have to be so general as to be of little practical significance to any one region in particular.

From the above discussion it is obvious that location theory cannot itself be applied directly to any one particular problem, for it has been shown that it fails to consider the conditions of demand, it fails to lay emphasis on the correct costs, its basic behavioural assumption is unrealistic and, perhaps above all, its aims, though theoretically noble, are too general to allow its conclusions to be of much practical significance. It will be seen in the ensuing chapters that various attempts have been made to take the essential ideas of location theory and, through the adoption of a more restricted and pragmatic form of analysis, to infuse a greater realism into its conclusions. It is to these that attention will now be turned.

Chapter Three

SYSTEM OF TOWNS APPROACH

In the last chapter location theory was found to be a highly abstract theory aimed at finding a general solution to a general problem. It purported to consider all possible factors which could influence industrial structure and then, in the best tradition of economic theory, it aimed at deducing a general equilibrium solution based upon those factors which were universally prevalent. In so doing, location theory adopted an essentially micro approach making the implicit assumption that if only one could understand the location of individual industries then the question of a town's industrial structure could look after itself. Implicit in the criticisms made of this theory were two important weaknesses. The first is that a general theory, through its obsession with what factors are always present, inevitably neglects the important question of the extent to which these may be over-ridden by purely local considerations. Secondly, in an attempt to define the whole, it is rarely feasible or advisable to study each component individually. This is a direct consequence of the first criticism and is due to the inevitable presence of a large number of cumulative deviations from the norm, thus giving the solution to the whole, as built from a study of its components, an unacceptable degree of variance. In addition, when it is further considered that location theory only embraced manufacturing activities, and only market and material orientated ones at that, it can be appreciated that as an attempt to explain a town's industrial structure the micro approach is inappropriate. As a result of these weaknesses a macro rather than micro approach should be adopted, which in practice means that towns should replace industries as the basic unit of analysis.

This is unfortunate for had location theory worked it would have provided the ideal solution to predicting a town or region's industrial structure; but its aim was too ambitious and its scope too wide for it to be of any practical value. The subsequent writings with which this and the next chapter are concerned owe much to the ideas of location theory, but at the same time represent attempts at overcoming its basic weaknesses.

region here treated as system of towns

In order to analyse the industrial structure of a town one may either study the town in isolation within a strictly economic framework or one may consider the town in the context of its relationship with those around it. The former approach will be dealt with in the next chapter, whilst the latter forms the subject matter for this chapter. Both avoid some of location theory's difficulties through the adoption of a less sophisticated aim, for they are not concerned with deducing general factors which influence towns, but rather with analysing towns, or systems of towns, as they are found in reality. Their aim is consequently more practical but less ambitious than that of location theory.

To turn attention specifically to the system of towns approach, the rationale for this lies in the fact that, from an economic point of view, towns show certain similarities in their industrial structures which permit them to be grouped. One branch of this approach, the statistical analysis method, goes no further than this. However the hierarchy of towns approach considers the question of whether groups may be arranged in the form of an hierarchy, so that the average industrial structure of the groups ~~bear~~ bears a definite relationship to each other. In either case the relevance to the thesis is obvious, for if towns can be so identified and categorised, then the variance in any one town's industrial structure can be immediately reduced through an understanding of which group it falls into.

HIERARCHY OF TOWNS

As stated previously, much of the writings with which this thesis is concerned find their inspiration in the workings of location theory and nowhere is this more true than in the development of the Hierarchy of Towns approach. This relies entirely upon central place theory, which itself is a direct descendant of location theory.

It is fair to consider both Lösch (32) and Christaller (7) as the two founders of central place theory. Although Weber (63) had previously developed a crude notion of an hierarchy it was left to Lösch to expound fully the way in which towns would be juxtaposed to each other. Centres of population (central places) would be arranged, it will be remembered, in a system of nets, the largest central place being surrounded by smaller ones and these in turn being surrounded by ones smaller than that. In this analysis Lösch is the first to develop in any significant manner the notion of an hierarchical order and, in so doing, lays valuable emphasis upon the inter-relationships between towns. Nevertheless it was left to Christaller to take these highly abstract ideas and to give them any practical expression. He did this by first dividing goods and services into two categories. On the one hand there are those goods the productions of which are geographically dispersed according to the location of their natural inputs, whilst, on the other hand, there are those central goods and services whose production consists of assembling the above dispersed goods in one central place. Christaller then takes up Lösch's notion of an hierarchy by sub-classifying these central goods according to the minimum populations required to support them. Then, given the implicit assumption of indivisibility in the minimum population required, an hierarchy of central places will develop according to the number and range of central goods produced at that point. For one central place to be on a higher order it is a necessary condition that it produce more central goods than the one below it, and it is a necessary and sufficient condition that it produce all those goods found in the order immediately below it, plus one more. Christaller identified different types of central places and these correspond, according to Ullman (60), to market hamlet, township centre, county seat, district city, small state capital, provincial head city and regional capital city.

But in addition to this notion of an hierarchical ordering, central place theory also includes the idea that these central places will be spaced regularly, with the distances separating any pair of settlements of like size increasing as one moves down the hierarchical ordering. Hence the rigidity of the whole framework will depend upon the strength of the assumptions made about the influence of distance. Christaller adopted one extreme position by maintaining that the above distance between similar central places would increase by the power $\sqrt{3}$ as one proceeded from one order to another. Perhaps the weakest, and by implication the most reasonable, interpretation of central place theory is that adopted by Bogue (4), who states:

> 'as the distance from the metropolis increases, the number of persons per square mile of land decreases. With increasing distance, each square mile of land area supports steadily decreasing average amounts of retail trade services, wholesale trade and manufacturing activities'.

Thus it is possible to deduce from Lösch's concept of an ideal market the idea of an hierarchy of cities and also to show that this system should have a precise spatial pattern. But all this is pure deductive reasoning and is not, in itself, a sufficient base upon which to build an explanation of industrial structure. What is required is evidence to show that this system manifests itself in reality. For this to be so it is necessary to show not only that there is evidence of such an hierarchy, but also that such an hierarchy is capable of being interpreted in meaningful terms. It would for instance not be helpful for our purposes to find that an hierarchy existed, but that the reason for its existence defied all reasonable explanations, for then one could legitimately suspect that its formation was the result of coincidence. The confusing fact about central place theory is that it neither passes nor fails the above two criteria.

To turn to the reasoning first, one would expect that an hierarchy of some form or another would develop. Each city will have some market-orientated activities (central goods) whose market area will be spatially defined. These cities will have different sizes since size depends partly upon the number of activities located within it and economies of scale dictate that not all activities can be produced in each place. In addition, economies of scale will have varying significance for different commodities and where these are the most dominant the commodity produced will tend to be a national good. Hence it might be expected that, for a variety of reasons, one city will capture the largest amount of these national activities. This would make it the first rank city, and other lower cities would be placed according to the range of these central goods. All this is common sense, but unfortunately central place theory goes one step further for it implies a statistically regular hierarchy of central places. Common sense tells one that towns may be ranked, but it tells nothing about their precise relationship vis-à-vis their hierarchical ordering. In other words within the framework depicted by Majumdar (34), reasoning would tell one that the hierarchy would be ordinal, whilst central place theory implies that the hierarchy is measurable up to a linear transformation (neo-cardinal).

Although one would perhaps not wish to go the whole way with central place theory, nevertheless it could never be denied that a priori the essential notion of an hierarchy of towns makes sense. But is such an hierarchy to be found in practice? There have been innumerable attempts at answering this question and for this brief survey it will be possible only to mention a few major works in each class. Since the first step in any such study must be the adoption of criteria by which to rank the constituent towns, it will be found convenient to categorise the various attempts according to the criteria used. Central place theory implied, strictly, that towns should be ranked according to population size and this led naturally to population being taken as one of the criteria. But the rationale of the central place scheme basically rests upon the notion of an hierarchy of functions between towns, the implicit assumption being that the greater the number of functions performed by any one town the greater will be its population. So an alternative criterion is not that of population, but rather that of what determines the size of this population, viz. the functions performed by the towns. So a survey of the evidence supporting the existence of a hierarchy of towns can be divided into those taking population as a criterion and those taking the function performed by towns.

Size Approach

To take the population studies first, the main contributors here are those studies based upon an application of the Pareto curve. This curve:

$$rP^q = K$$

relates the size (population) of a town (P) to the number of towns sized P or more (r), K being the number of towns in that class. This formula has been held to have widespread validity. Zipf (65) found that the exponent was nearly unity when applied to the 140 metropolis districts which had populations of more than 50,000 in the U.S. Further studies, especially those by Allen (1), have shown that on the whole data for a number of different years, countries and range of sizes of towns all give an acceptable fit to the Pareto formula. Though it would appear, according to Clark (8), that this rule becomes less acceptable the more industrialised a country becomes.

What the persistent (some would say stubborn) ability of the formula to fit the facts illustrates is that towns show a systematic and regular hierarchy. Further, if the exponent is unity (for which there is considerable evidence) then the formula becomes the familiar rank-size rule, which means that the size of a given community can be expressed as the quotient of the size of the largest community divided by the rank of the given community.

But what of its real significance? Critics have pointed to three weaknesses of this rule. The first is rather technical, the other two more profound. The first is associated with Hoover (24) who raises the question of delimitation of cities. The rule apparently works satisfactorily for the U.S. so long as one starts with New York, but it would fail to hold good should one have chosen to start further down the hierarchy. Thus for the rule to be valid the system of cities must be defined carefully at the outset. Duncan (12) has summarised the second weakness of the rule when he states that:

> 'although the Pareto distribution is compatible with the central-place scheme, an empirical fit of the Pareto curve hardly validates central place theory in detail' (p.55).

The reasons for this being so are two-fold. First the rule gives a smooth descending curve whilst central place theory would imply a step-wise function, each step being associated with a discrete grouping of functions. Secondly Stewart (52) emphasises that the value of K will not be universally constant, but will alter from region to region. This is due, principally, to the presence of more sophisticated services in the higher order of towns, the demand for which will alter with a region's income. What both these objections really mean is that although the Pareto distribution would initially appear to support the central-place schema, through its delineation of an hierarchy of cities, it fails to do so in detail since in the last analysis the hierarchy is explicable in terms other than that of central place theory.

Any connection that still may have been thought to exist between the two is finally severed by the third criticism of the Pareto curve, namely that of the already raised fundamental question of theoretical justification. Here the apparent universality of the principle is both its strength and its weakness, for if it applies everywhere, then surely the same behavioural laws must be in operation everywhere and, by implication, discernible. If, as one suspects, the law is incapable of theoretical interpretation then it cannot be held to support any particular theory.

So, to summarise, the rule shows that there is a greater degree of regularity in the ranking of towns and cities than one would originally suspect, and, as such, it supports the notion of an hierarchy of towns. But since such an hierarchy is not related to the theoretical scheme of central place theory, any support which the rule may be thought to give to that theory per se must be discounted.

One second approach which adopts population as a criterion must be mentioned briefly. Duncan (12) used a size-classification in ordering the towns of the U.S. For each class interval he determined the average industrial structure, for he presumed (reasonably) that if there is an hierarchy of cities, then it would be likely to manifest itself in a towns industrial structure. It follows, therefore, that the further apart the size class intervals are, the greater will be their dissimilarity in industrial structure. Again evidence for an hierarchy was found. Further, its formation this time was more explicable in terms of central place theory, for it was found to be due principally to the tendency of extractive, processing and local services industries to relate inversely to size and the complementary tendency of fabricating and non-local services to vary directly with size. Central place theory would certainly imply that the more sophisticated fabricating and non-local service industries would vary directly with size. This is an important result which will be returned to later, but in this context it should be noted that the presence of an hierarchy had to be inferred since the size of class intervals was assumed beforehand.

Function Approach

The second group of studies tries to delineate an hierarchy in terms of the functions performed by the towns. These may be defined either by looking at that town and the others around it, or by investigating the town in isolation and categorising the functions it performs. In either case the procedure is the same, viz. first postulate criteria by which to rank towns and, secondly, on the basis of assumptions made about the qualifying functions for each class, assign towns to different levels in the hierarchy. For the first approach a single criterion is adopted whilst, typically, the second relies upon a battery of statistical indicators.

The first approach tries to define the relationship between a town and the rest of the system in terms of bus routes into and out of that town. The classical work in this field was done by Carruthers (6).

First he took as a criterion for a centre the presence of at least one bus service per week which exclusively served a population smaller than itself, then he plotted on a graph the number of such services per centre and from the resulting pattern he purported to identify (subjectively) an hierarchy in the importance of those centres. For instance from the above map it would be obvious that Reading is more important as a centre than nearby Maidenhead, which in turn is more important than Wokingham. In this manner Carruthers identifies 1, 2A and 2B, 3A, 3B and 3C and 4A, 4B and 4C orders. The procedure has been repeated for Scotland (though only in connection with the lowest order) by Fleming and Green (15). Further Taaffe (56) in a study of U.S. centres substituted aircraft for bus services (which meant that he had to exclude the smaller centres).

The difficulty with these approaches is, first, that although bus travel is an obvious measure of centrality in aggregate, it need not be so for each and every centre in particular. One would at least like to see the adoption of some further criteria so as to ensure against the presence of unusual characteristics in any one town's bus services. But nevertheless the most profound difficulty lies in the assumptions defining each category. At least in the studies to be described below there is some attempt to make these objective, whereas in this study it really is a question of unfettered personal judgement. For how else could it be maintained, on the basis of bus data alone, that it is possible to distinguish between Cardiff as a 2A centre and Liverpool as a 2B centre?

Whilst the first approach focused attention upon placing a town in a particular category by analysing its relationship with other towns, the second focuses upon the functions performed by the towns themselves.

There have been two types of studies concerned with the functions of towns. The first will be mentioned only to be dismissed. This, the so-called functional specialisation approach, consists of identifying the dominating (single) function of a city. The criteria for classification may be subjective Harris (21) or objective Nelson (39), but both types suffer from the weakness that they do not attempt to arrange towns in any form of an hierarchy. Also, since few towns can be identified on the basis of a single function, it is not surprising to find that the classification is not mutually exclusive. For instance it often occurs that a 'finance' town may also be a 'manufacturing' one as well. These studies lack any rigorous framework and are never more than mere descriptive devices. The second approach manages to avoid these criticisms for it explicitly sets out to place towns into an hierarchy. Some of these studies have simplified the problem by considering only one end of the spectrum. Thus Brush (5) considers only hamlets, villages and towns whilst Green (19) emphasises distinctions between metropolis and provincial capitals. The example to be quoted here is that of Smailes (51) who considers the whole spectrum. He divided towns up into London, major cities (either regional capitals or provincial cities of great importance), cities, other centres (minor cities or major towns), towns, sub-towns and urban villages. He uses as his criteria a whole host of statistical indicators (viz. employment in insurance, wholesale trade, bank clearances, Bank of England branch, Stock exchange, etc.). Then his various categories are defined through assumptions made about the qualifying function for each class. Thus for a centre to be a town it must possess three to four banks, two cinemas, a local newspaper and a secondary school and/or hospital. The distinction between major cities and cities turns upon the fact that the former have a general regional importance, whilst the latter may be important, but only on an ad hoc basis. With the aid of the above criteria and assumptions, Smailes is able to delineate an hierarchy of towns in England and Wales.

The two approaches to the problem of identifying an urban hierarchy suffer from considerable weaknesses associated with both the adoption of their criteria and their assumptions. Both these weaknesses have the same practical outcome of making the hierarchy identified vary with each study. The main problem with the criteria is that they are usually too subjective. Smailes's criteria for defining the category of towns may be reasonable in themselves, but they do not preclude the possibility that another equally reasonable list of criteria would not lead to a different set of towns. Yet really this is a criticism more about the nature of the problem itself rather than about these attempts in particular. One could avoid the above criticism by adopting a single criterion for classifying towns. This is in fact what the first category did with their bus and aircraft studies, whilst in the second category Siddal (50) used wholesale/retail ratios as a single criterion. Yet such a procedure hardly makes the resultant hierarchy any more meaningful for it presumes (as already stated) that towns can be uniquely identified in terms of one attribute (function). The real difficulty is that each town is characterised by different functions and it follows therefore that the very nature of the problem is such that the choice of what functions to adopt must always be subjective. It is for this reason therefore that it is not surprising to find that the final hierarchy discovered varies between each study.

The second weakness of this approach lies in the fact that since the constituent elements in the hierarchy depend exclusively upon the assumptions made about what functions are to be associated with each class, the choice of these elements is in effect arbitrary. This has two consequences. First, in an attempt to get a few meaningful classes, most authors adopt wide assumptions, which has the immediate result of making these classes capable of further sub-division. By way of example, Smailes's 'major cities' can be sub-divided into regional capitals and provincial cities and Carruthers' third order centres are further divided into 3A, 3B and 3C centres. From this it would appear that the hierarchy is rather more homogeneous than one would be led to believe and since the whole usefulness of an hierarchy lies in an apparent degree of discreteness, the principle must be somewhat undermined by these considerations. The second consequence is the corollary to the above, namely that since the divisions are chosen arbitrarily, there is a tendency to assume the presence of a division in the first place and then to prove its existence.

To a certain extent these difficulties with the assumptions have been overcome by Berry and Garrison (3), who have shown that assumptions can be dispensed with through the generation of categories statistically. They identified 63 central functions and ranked towns according to their possession of these functions and then tested the sample for grouping (where a group is defined so that all its members are closer to each other than to any non-member). On the basis of a purely statistical procedure they identified an hierarchy, thus avoiding the need to make any assumptions at all. Nevertheless it still has to be shown that the resultant hierarchy is any more meaningful than others.

The third and most damaging weakness of these approaches is the natural outcome of the above remarks on criteria and assumptions. Central place theory does not imply that each country will exhibit the same hierarchy, but what it does imply is that there will be one hierarchy per region, however indistinct that may be. It implies, in other words, the presence of a natural hierarchy for each region. But these studies show that there is a different hierarchy for each package of criteria and assumptions. For instance in Carruthers' study both Southampton and Sheffield are 3A centres, but Smailes classifies them as 2B centres, whilst both agree that Plymouth and Norwich are 2B centres. This would imply that either the methodology is suspect or there is in fact no unique and natural hierarchy for each region.

Urban Hierarchy: Conclusions

From the essentially a priori reasoning of central place theory comes the idea that towns are not arranged in a homogeneous spectrum, but rather that they are to be found in groupings, arranged in the form of a strict hierarchy. Although the rationale for this lies in the greater range and sophistication of functions performed by the larger towns, it has been assumed that this hierarchy could be expressed in terms of population size. However, in an attempt to prove or disprove the conclusions of central place theory, studies have adopted both population and function as a criterion for delineating urban systems. The difficult and confusing part about urban hierarchy is that central place theory is shown to be neither completely correct nor incorrect. Basically this theory states (1) that there will be an hierarchy of towns either by function (industrial structure) or size, (2) that this need not be entirely distinct and (3) that there is only one (natural) hierarchy per region. All the various studies point to an hierarchy in some form or another, for one of the main findings common to all is the presence of relatively few classes of towns, each with a relatively large functional gradation. As would be implied by central place theory, the presence of only a few classes can be interpreted in terms of indivisibility of units of production and complementarity in space between different products and services. But beyond this the studies fail to confirm the conclusions of central place theory, for the simple reason that the various approaches fail to identify one single and dominant hierarchy. This means, first, that any compromise hierarchy that could be constructed out of the individual hierarchies would be too ill-defined to be compatible with central place theory. It is true that this theory permits a certain indistinctness due to the heterogeneity of natural resources and environmental conditions, but what is present here is far and away above the permitted level. Secondly, these analyses show that there can be no such construct as a 'natural' hierarchy, but rather that there is one for each behavioural viewpoint. Surely it is this that is the crux of the whole matter, for clearly each economic activity is associated with an independent and separate hierarchy. There may be every reason to believe that there will be an hierarchy in, say, manufacturing and also one in administration, but there is no reason to presume that they will coincide (witness the different levels pertaining to Edinburgh and Glasgow in two such hierarchies). It may be true that Lösch defined the most important one when he considered function, but when it is realised that this cannot be the sole hierarchy the conclusions of the various analyses described in this chapter immediately become more intelligible. Clearly there will exist several significant hierarchies and also any attempt to construct a natural or generic one must lead to indistinctness.

Finally, what are the conclusions that can be taken from the above in connection with a study of a town's industrial structure? This is a point which will be taken up in much greater detail later on, but it should be noted that although central place doctrines cannot be applied wholesale, there is still ample evidence of a non-homogeneous spectrum of functions. Further, any narrowing of the field of inquiry will lead to this effect becoming more pronounced. In other words to understand the causes of differences in towns' industrial structures or to interpret the findings of empirical analysis, it will always be necessary to consider a town's position in relation to any hierarchy that may be present.

STATISTICAL ANALYSIS OF TOWNS

In the introduction to this chapter it was stated that there were two approaches to analysing a system of towns. The first, that of the hierarchy of towns, having been dealt with it is now time to turn attention to the second, that of the Statistical Analysis of Towns. This approach, which is entirely based upon the work of Moser and Scott (37), differs from the urban hierarchy analyses in both its methodology and its aim. The latter tried to categorise towns in a special way, i.e. they tried to rank towns according to their position in an hierarchy. Moser and Scott, on the other hand, are only concerned with classifying towns, and in this respect their aim coincides with that of Harris (21) and Nelson (39) though, as will be shown later, without suffering the latter's deficiencies. Their difference in methodology is two-fold. First, urban hierarchy, in its selection of criteria, relied upon *a priori* reasoning to keep the battery of statistical indices within manageable proportions. Moser and Scott, through the adoption of powerful statistical tools, are not subjected to this constraint and can handle any number of criteria. Secondly, the urban hierarchy approach required the adoption of assumptions to delineate the hierarchy's structure, whilst Moser and Scott allow the categories to be chosen by statistical means (though they do not go quite as far as Berry and Garrison (3) in this respect).

Moser and Scott take 57 variables relating to all aspects of a town's life ranging from its health to its voting habits. Thus, although 10 economic variables are included (basically relating to the employment structure of towns), the analysis is concerned with the whole spectrum of a town's character and is not therefore exclusively economic in scope. Next these statistics are compiled for each of the 157 towns with populations in excess of 50,000 in England and Wales, and from these a correlation matrix is constructed showing the product moment correlation co-efficients between each pair of variables. The matrix itself is sufficient to show that most of the variables are inter-related and their next step, with the aid of component analysis, is to discover how much of the variance between towns can be explained by a smaller number of independent variates (components). Component analysis ensures that the first component accounts for the greatest possible proportion of the total variance, whilst the second, which is completely uncorrelated with the first, then accounts for the next greatest part of the variance. In this study the amount of variance explained is relatively low for which the wide diversity in both the variables and the towns is directly responsible. Nevertheless the first four components account for 30%, 13%, 10% and 7% respectively, i.e. a total of 60%. Moser and Scott feel confidence in these results for essentially the same conclusions are found if logarithms are taken for the variables or if rank-order correlation co-efficients are used. These components, which are artefacts of the technique, may be interpreted in the light of those prime variables which are most closely correlated with them. On this basis Moser and Scott associate their four components with social class differentiation, population growth 1931-1951, population growth 1951-1958 and overcrowding in that order. If the towns are now plotted in a 4-dimensional space, taking as co-ordinates their values for each of the four components, then it is possible to group these towns statistically. This requires the two assumptions of, first, that a town will be in one group if its distance from the centre is less than its distance from other centres and, secondly, that a group shall be considered to be formed only when it contains ten towns. Applying these assumptions to their data, Moser and

Scott find that it takes 14 groups to accommodate all their towns (with the exception of London and Huyton which are too extreme to be classified).

This then is their analysis, but what of its significance for the purposes of this thesis? Its basic contribution lies in the fact that Moser and Scott manage to categorise towns using the widest possible criteria with the minimum number of assumptions and that the resultant category is based upon those which are the most fundamental to towns. This represents a considerable advance over the urban hierarchy approach, for these criteria had to be adopted on an a priori basis so that it could never be certain that the fundamental ones had been chosen. In the sense that Moser and Scott find all the components to be non-economic it is not surprising that all the attempts to identify hierarchies in purely economic terms should lead to different results, for obviously none of these factors are basic to towns.

The disadvantages of this technique are, first, that it is purely descriptive since it does not attempt to place towns in an hierarchy, and as such it lacks any behavioural content, though unlike Harris (21) and Nelson (39) the classification is mutually exclusive. Secondly, although the study goes a long way towards making the selection of categories objective, it never quite reaches this since it still requires the assumption of what is the minimum number of towns necessary for a group to be formed. In other words, it lacks the complete objectivity of Berry and Garrison (3). Further, since the deletion of data prior to 1951 would obviously alter the results, the classification must be considered subjective since it depends very much upon what data is fed into it. Finally, it could never be maintained that this category was in any way absolute since the four components only encompass 60% of the variance.

CONCLUSIONS TO SYSTEMS OF TOWNS APPROACH

This chapter has had a two-fold aim. The first to expound and test the ideas of central place theory, the second to describe the evidence available for non-homogeneity in the economic spectrum of towns. As will be seen, both these ideas form a vital background to the empirical work to be described later.

It is clear that central place theory is invalid in its rigid form. There can be no single or unique hierarchy, but rather several, the most significant of which relate to the basic characteristics of towns. Moreover the section on the statistical analysis approach showed quite clearly that the characteristics which are fundamental to towns are not even the economic ones. But central place theory does state that the range of goods offered will vary with population and, although towns cannot be assigned to any rigid hierarchy, there are considerable grounds for believing that the essential notion behind central place theory is still valid. There is, first, the whole host of evidence illustrating that towns can be grouped or categorised. Secondly, although these groups may not be capable of any hierarchical juxtaposition when viewed on a national scale, it nevertheless remains clear that within a reasonably small area such an hierarchy exists. These two ideas are important for it means that the functions performed by towns cannot be interpreted without considering both what category the town falls into and what is its position within the hierarchy associated with its immediate area. Quite how significant, though, will be a matter for further empirical investigation.

<u>Chapter Four</u>

BORROWINGS FROM MACRO-ECONOMIC THEORY

Viewing the last chapter in its broadest context, it was concerned with studying how demographic and geographic factors could influence industrial structure. More precisely the chapter tried to discover whether the industrial structures of towns show any hierarchical ordering, or any other form of grouping, which is interpretable in terms of demographic variables. Although the rationale for such an inquiry was based upon the findings of the location theory, there was no attempt to apply economic theory <u>per se</u> to the problems of industrial structure. It is the intention of this chapter to redress this imbalance by considering how economic theory can be applied to the problems of industrial structure.

There is a wide range of choice of economic theory that can be applied to any particular regional problem, yet it is not the intention of this chapter to review the whole body of this theory but rather to limit the survey to those parts which are relevant to the problem at hand. Implicit throughout the whole of regional economic theory are the difficulties associated with the lack of proper regional statistics, and this problem is made even more acute in this thesis through the concentration upon towns rather than regions. This effectively reduces those theories which are appropriate to the problem at hand to the Sector and Regional Multiplier theories. These will be considered in detail in the first two sections, but in the last section some attempt will also be made to briefly outline the remaining theories and to justify their exclusion on the grounds of their inconsistency with the available data.

SECTOR THEORY

This theory finds its origin in the writings of Clark (8) and Fisher (14) in which they analyse the relationship between the level of a region's income and its industrial structure in terms of employment. Clark, in his study, conducted a survey of 33 countries and compared their industrial structures with their average real per capita income, both on a cross-section and a time series basis. From his empirical observations he concluded that:

> 'we find a very firmly established generalisation that a high average level of real income per head is always associated with a high proportion of the working population engaged in tertiary industries. Low real income per head is always associated with a low proportion of the working population engaged in tertiary production and a high percentage in primary' (p.7).

He defines primary industries as agriculture, fishing and forestry, secondary industries as manufacturing, mining and building, whilst tertiary industries are the remainder. Further

> 'The reasons for this growth in the relative number of tertiary producers must largely be sought on the demand side. As income rises .... the demand for such services increases and, being non-transportable, they must be supplied by workers within the country concerned' (p.7).

These simple but highly suggestive ideas have been taken up by others and considerably refined. First it was realised that the shift in the labour force was due not only to the conditions of demand (viz. the higher income elasticities of tertiary products) but was also due to changes in the supply conditions, i.e. the greater scope for productivity increases in the primary sector vis-à-vis the tertiary sector. Secondly, these ideas were combined with Weber's notion of a stratum of economic activity to yield the 'Theory of Development Stages'. Here again one has an example of location theory's basic ideas being taken up by later writers, this time to be given a more sophisticated behavioural content. The various writers in this field, especially Hoover (23), by extending the sector theory have managed to depict the 'normal' sequence of a region's development through the following stages (1) a self-sufficient subsistence economy, (2) product specialisation in primary activities allied to some inter-regional trade, (3) the introduction of secondary industries, (4) the shift to more industrialisation based upon internal industrial linkages and finally (5) an advanced stage of economic development and specialisation in certain tertiary industries for export. Thirdly the tertiary sector has been disaggregated by Foote and Hall (16) into purely tertiary services (domestic and quasi-domestic), quaternary industries (transport, communications, commerce, finance and administration) and quinary industries (including medical care, education, research and recreation).

These, the essential ideas behind the sector theory, are obviously of considerable interest to this thesis. What Clark and Fisher have shown is that the higher the income of a region <u>ceteris paribus</u> the greater will be the size of the tertiary sector, so that in any attempt at explaining differences in industrial structure the respective levels of income must play an important part. But the question is how important a part? Although neither Clark nor Fisher offered any systematic analysis of the factors lying behind the differential shift in employment, this has not deterred some economists from going to one extreme by presuming the existence of a rigorous quantitative relationship between the level of income and the size of the tertiary sector. Admittedly on an initial reading of the data the circumstantial evidence for such a presumption is strong. Stigler (54) has reported that the share of the U.S. labour force engaged in services has increased from 20% in 1870 to 50% in 1950 and Kuznets (29) has estimated that consumer expenditures have increased more than fourteen-fold between the decades 1869-78 and 1919-28, whilst total consumer expenditures have increased slightly more than eleven-fold.

But it would be unrealistic to adopt such an extreme position, for when the data is considered in any real detail the explanatory power of the income variable is somewhat weakened. First, as Regan (44) has shown, the trend slowed down in the period 1929-1959. Although he would prefer to explain this principally in terms of a re-classification of industries, together with the tendency for the government to take over manufacturing activities, it is nevertheless clear that doubts must remain about the strength and consistency of this trend.

More importantly, doubts about the Clark-Fisher generalisation must arise from its theoretical justification in terms of differential income elasticities and labour productivities. The pitfalls of this analytical base have been summarised by Bauer and Yamey (2) and one can do no better than quote them at some length. They maintain that the generalisation

> 'is open to criticism on several independent grounds. First a substantial proportion of tertiary products are not luxuries with

> a relatively high income elasticity of demand; conversely, some products of primary and secondary production, possibly on a large scale in their aggregate, are such luxuries. Secondly, there may be large scale substitution of capital for labour in tertiary production in the course of economic progress. Thirdly, the concept of the income elasticity of demand applied to the whole economy raises problems of aggregation which render doubtful any universal proposition about changes in its average value in conditions of change and economic growth' (p. 748).

These three criticisms may be considered in turn. The first arises from the fact that tertiary production is a heterogeneous collection of different services. Some of these are income-elastic luxuries, but nevertheless still a large part of the tertiary sector is indispensable to any stage of economic development. Undoubtedly this is fair *a priori* reasoning, yet Bauer and Yamey fail to provide any evidence to support their contention that, as a result, income elasticities will show no significant variation between the sectors. However work conducted by Fuchs (18) does. To determine whether there is any systematic difference between the income elasticities of the three sectors is hazardous, chiefly because of an inability to measure the real output of services, together with the difficulty of dissociating employment shifts due to income changes from those due to urbanisation. These difficulties apart, in a period of rising real income a higher income elasticity in any one sector would imply a higher growth rate of real output for that sector, yet Fuchs manages to show that such was not the case for the period 1929 to 1963. Since Fuchs's empirical work would tend to support Bauer and Yamey's analytical reasoning, one must perforce reserve judgement upon the efficacy of the demand interpretation postulated by Clark and Fisher.

Granted that there has in fact been a differential shift towards the tertiary industry in this period, albeit at a slower rate, then it follows from simple accounting principles that if this change is not due to differences in income elasticities then it must be due to a greater rise of productivity in the primary and secondary sectors. This conflicts directly with Bauer and Yamey's second criticism, namely that there is scope for large scale capital substitution in the tertiary sector. Although again the facts must be allowed to settle the issue, one would perhaps initially wish to quarrel with their arguments. First Bauer and Yamey's examples of whole-scale substitution (at least for an advanced economy) are rather weak, resting basically upon an (unspecified) scope in domestic, laundry and repair services. That apart, they maintain that the sector theory would be valid

> 'only if it were legitimate to assume that labour and other productive resources were employed in tertiary production in fixed proportions' (p.749),

whereas in fact all that is required is that the scope for substitution be less in the tertiary sector than in the other two.

As far as the statistical evidence is concerned, this is far from being conclusive. Fuchs does show that the differential change in employment is explicable almost entirely in terms of changes in output per man. But interpreting these results is not easy. First of all the initial partition of changes in employment into changes in output and productivity has its limitations since the two are not independent. For instance, shifts in output can affect productivity through economies of scale and the concomitant stimulus to technological change. Secondly, the connection between 'labour productivity' as implied by Clark and Fisher and 'output per man' as measured by Fuchs is somewhat tenuous. There are considerable grounds for believing that the faster growth of output per man may be due at least in part to the greater number of hours worked in the secondary sector, along with the employment of a qualitatively superior labour force. Admittedly the end result may be the same, but the implications for the future are not since these could well be once and for all changes. Nonetheless, Fuchs is still of the opinion that part at least of the differential growth of employment in the tertiary sector is due to the substitution of capital for labour in response to technological changes within the secondary sector. To the extent that this is likely to be a continuous process, then one is perhaps justified in maintaining at least some faith in the Clark-Fisher hypothesis.

The third criticism of Bauer and Yamey, as applied to a developed economy, is a consequence of both their first criticism and of the distribution of income. If not all the luxuries are exclusive to the tertiary sector, then the greater rise in productivity in the secondary sector will lead to a shift in demand for luxuries away from those of the tertiary sector towards those of the secondary sector. Secondly, even if the majority of the luxuries are to be found in the tertiary sector, then the growth in demand for these vis-à-vis the less sophisticated products of the secondary sector will depend upon the distribution of income. Yet, in assessing the impact of both these criticisms, all that can be said is that a priori they are valid, but again there is no evidence given (or available) by which to judge their practical importance.

Conclusions to Sector Theory

It has been shown that the Clark-Fisher hypothesis is open to several criticisms which aim basically not so much at refuting the existence of a shift of labour into the tertiary sector, but rather at questioning whether their findings have any behavioural content. The analytical basis of the Clark-Fisher hypothesis rests upon the key assumptions of a systematic variation in income elasticities and the growth of labour productivity as between the three sectors. To the extent that evidence can be adduced not only to support the latter, but also to suggest that it is due to a trend which is likely to persist into the future, then, in this respect, the criticisms must fall short of their aim.

On the other hand, the presence of variations in income elasticities is certainly not proven, but this only serves to raise the further question of why the trend fails to manifest itself. It is here that one runs into the time-honoured aggregation problem which invariably arises when statements aiming at simplifying generalisations are made. The more one simplifies and aggregates the more useful become the conclusions but, at the same time, the more imprecise become their practical effects. It is without doubt that Clark and Fisher are guilty of oversimplifying the situation, for clearly not all tertiary products are luxuries, neither are all luxuries found only in the tertiary sector. One would certainly wish to agree with Rottenberg (46) that Bauer and Yamey:

> 'do strike effective blows against the Clark-Fisher thesis that at low incomes purchasers will necessarily prefer material things to non-material things'. (p.168)

but one would suspect, unlike Rottenberg, that this is not because the fundamental premise is incorrect, but rather because it is concealed by too aggregative an approach. One must therefore agree with Perloff's (41) conclusions based upon a survey of 48 states in the U.S.

> 'That one cannot conclude from such evidence that the association between the level of income and employment patterns is not analytically significant. What suggests itself is that there <u>is</u> a significant relationship between income levels and industry (employment) structure, but that this relationship is not best analysed by the three-way classification employed by Clark and others. The evidence suggests that, for the states of the United States at least, Clark's groupings are probably too aggregative, hiding highly important variations within each of the three groups.' (p.165).

In the last analysis, the criticisms of Bauer and Yamey illustrate why the sector theory in practice must be treated with caution, but since they fail to actually disprove the sector theory's analytical framework, one must allow the essential notion of an association between income and industrial structure to stand.

REGIONAL MULTIPLIER THEORY

In the last section the deterministic approach of the sector theory led to an explanation of industrial structure cast purely in terms of the level of the region's income which was assumed to be given. The regional multiplier theory requires no such assumption, but rather tries to analyse differences in industrial structure through a study of the determinants of a region's income. In order to do this it has to extend the concept of a region found in the sector theory. First, Clark and Fisher were concerned with the region in isolation, whereas the relationship between the region and the rest of the world forms a principal part of the multiplier studies. In other words, in the regional multiplier approach, the region must be viewed in a wider context. Furthermore this approach is firmly based upon the dichotomy between that part of the economy which earns and that part in which the income is spent, and this is in sharp contrast with the concept of a region as an indivisible unit adopted by the sector theory. Perhaps partly as a result of these differences in approach, the two theories are usually considered as being alternative frameworks of analysis [see especially Thomas (57)]. However, to a certain extent the regional multiplier analysis merely takes the sector theory one step further through its consideration of the determinants of a region's income. Nevertheless it is still true to consider the theories antithetical due to the wide disparity in their connections between income and industrial structure. The sector theory, through differential income elasticities, looked to the consumer for this relationship, whilst the multiplier theory is concerned with the purely empirical connection of the extent to which one sector can be supported by another.

However the two theories are not entirely devoid of similarity, at least to the extent that they share the same degree of abstraction and, therefore, the same weaknesses. Implicit in the criticism made of the sector theory was that it was only partial and, more importantly, that whilst it was a highly suggestive generalisation, it was nonetheless too aggregative for any analysis in depth. As will be shown later, one will have reason to agree with Perloff (42) that this applies to the regional multiplier approach as well.

Regional multiplier analysis, as might be inferred from its name, is similar to the normal Keynesian multiplier analysis in which changes in one sector are transmitted to and diffused through other sectors. Although the regional multiplier analysis conceptually can be designed tohandle any number of variables, according to Isard (27) the most comprehensive analysis that is still capable of giving useful empirical results is that of interregional input-output. This analysis depicts not only the inter-relationships between the region's own industries, but also those between these industries and other sectors. However, for the purposes of this thesis, the most important application of the regional multiplier (which is also the simplest) is that of economic-base theory, which itself might be considered a crude form of input-output analysis.

The economic base theory distinguishes between those industries which export their output out of the region (basic industries) and those whose output is sold principally within the region. The distinction is based upon the popular premise that the existence and growth of a region is determined by the goods it produces locally but sells outside the region, for it is the income from these 'basic' activities that pays for the region's imports and supports its non-basic sector. Consequently an increase in the output of the basic sector will lead to an increase in the size of the non-basic sector, the actual extent of the increase being determined by the regional multiplier. The latter is simply the ratio of the basic to non-basic industries in the region, as delineated by historical analysis. Its actual size will depend principally upon the community's marginal propensity to consume and to import. The exporting industries will consist of the bulk of the region's secondary production, together with those services which it undertakes for other regions, viz. insurance, banking, finance, etc. The non-basic industries will consist chiefly of the purely local service industries (retailing, local administration, perhaps transport), some secondary activities (quarrying, production of bricks, etc.) and the construction industry.

Just as the sector theory held considerable promise in determining the industrial structure of regions, at least in an aggregative manner, so does the economic base concept and it is possible, on the basis of this theory, to interpret differences in industrial structure in terms of different sizes of the basic sector. But also like the sector theory, the economic base theory is too simplified to be used as the chief explanatory variable. In order to appreciate why one must study the deficiencies and weaknesses of this approach. Basically these fall under the headings of technical and conceptual problems.

To consider the technical first, these are associated with the size of the base, the unit of measurement, the variability in the multiplier and the size of the area. The first, which is perhaps the foremost difficulty, concerns the problem of identifying the basic and non-basic components. This involves answering questions of the following nature. Does a firm exporting, say, 60% of its output qualify as a basic industry, and if it does at what percentage would it fail? How does one treat vertical integration? For instance does an industry exporting cement but owning its own quarry qualify entirely as a basic industry? Finally, if coal is mined locally, and sold to a local steel plant exporting finished steel, is the coal to be considered a basic or a non-basic activity? These are difficulties stemming principally from the inter-relatedness of modern industry and pose peculiar problems which can only be solved through either the use of more sophisticated tests (viz. interregional input-output analysis) or the making of some pretty arbitrary assumptions. But even before reaching this stage, the question of actually how to identify the exporting industries has to be solved. A separate survey of each industry would give the most accurate results, but it is obviously so impractical that researchers have had to turn to two other methods. Hildebrand and Mace (22) classified industries by their location quotients, taking any industry with a location quotient greater than unity as being an exporting industry. But unfortunately such a procedure necessitates the assumption of a coincidence between local and national consumption patterns, together with the assumption that all local demands are entirely met by local production. The second is to estimate the base through comparative cost techniques, i.e. by analysing each industry to see whether its costs are lower than elsewhere (if so it is deemed to be an exporting industry). But for small regions the variation is not likely to be all that great and for a study where more than one region is contemplated the method is obviously just as impractical as the firm-by-firm approach.

The second technical difficulty is associated with taking employment as the unit of measurement - (which is invariably adopted due to the lack of any other data). But this measure fails to catch the full significance of an expansion of employment in a high wage vis-à-vis a low wage industry; and neither does it reflect the expansionary effects resulting from an increase in labour productivity whilst the labour force remains constant. In other words, conceptually the method requires to be cast in an income framework, whilst in practice employment figures only are available and this raises certain technical difficulties. Researchers have tried to overcome this by weighting the employment figures with payroll or unearned income data, but it is still not clear whether the extra work involved is really worth while.

With regard to the third technical difficulty Friedly (17) has pointed out that the economic base concept is valid only to the extent that the multiplier is constant. As already stated, the value of the multiplier is determined principally by the average marginal propensity to consume and the marginal propensity to import. Over time both these may alter, though in which direction it is difficult to forecast. Take, for instance, a rise in income. One would expect this to lead to a fall in the marginal propensity to consume but to a rise in the marginal propensity to import as expenditures are shifted more towards services and other non-essential goods. Tiebout (58), on the basis of indirect evidence, feels that the overall propensity to spend locally will be constant, but as Friedly has shown this may be true for an isolated area (far removed from the effects of supermarkets etc.), but for an area near a sophisticated shopping-centre this may not be so. Friedly found that for a medium-sized suburban community in Los Angeles there was a similar decline in the multiplier, which was apparently due to an influx of new higher-income residents who did not have the traditional shopping habits of the older community.

Whether or not the multiplier will increase or decrease is an open question but all that one can say is that a priori it is likely to change. This is so not only for the above reasons, but also because, first, at any point of time, as a result of lags in adjustment, there is unlikely to be an equilibrium position, causing the multiplier to alter in response to past changes. Secondly the very essence of economic development involves the modification of past relationships, tastes, preferences etc. and therefore the multiplier, which is based upon these ties, is bound to alter with a rising real income.

The fourth technical difficulty involves the bothersome problem of choosing the correct geographical area. The analysis should be based upon a recognisable unit of population and the area should be so chosen as to include those manufacturing industries associated with this population unit, together with the local services supported by these industries. That this is awkward to do in practice is probably obvious, though a discussion of quite to what extent will have to await the empirical section of this thesis.

These then are the technical difficulties associated with the economic-base concept, but none of the criteria are fatal to the principle and many of them can be solved, or at least mitigated, in the course of further research. But this still leaves the question of the conceptual difficulties and one has to ask to what extent do these fundamentally undermine the economic base theory?

The main conceptual difficulty revolves round the controversy of whether the method is to be considered a short or a long-run device. The proponents of these two views are Tiebout (58) and North (40) respectively. The issue basically rests upon the extent to which exports lead, or principally determine, a region's economic growth. To the extent that they do not, then the concept must be partial and, therefore, only of short-run application since in the long run other influences will not remain constant. Tiebout sees the regional multiplier as a valuable aid to understanding income determination, but by no means considers it the sole determinant. In order to understand this viewpoint one must consider the two sectors in the region in the wider framework of exogenous and endogenous sectors. The income in the region is the total of the income generated in these two sectors. The income generated in the exogenous sector is outwith the control of the region. The recipients of this income spend it (partly) within the region and the generation of income that results from satisfying this demand identifies the endogenous sector.

Thus, since the level of activity in the endogenous sector depends upon the level of demand generated by the exogenous sector, it is legitimate to consider the latter as the basic sector. To the extent that the export sector forms a vital part of the exogenous sector, Tiebout would agree that it is important in the short run, and this is further reinforced by the fact that the exports will be the largest and most volatile part of this sector. Yet exports are by no means the sole

Sel. Sector's income

× constituents of this ~~sector~~, which must also include the level of investment, the degree of non-local government spending within the region etc. It follows from this that it would be perfectly possible for growth to take place through a change in factors other than exports, viz. a change of investment making its effect felt through the accelerator.

North, on the other hand, visualises a more deterministic course of economic development by assuming that at the start of economic growth an export sector comes into being and, as a result, dependent residentiary industries are established. In the long run it is only changes in the export sector that can initiate growth. But apart from ignoring all the other factors in the exogenous sector, such a theory would fail to allow for the establishment of industries which were not dependent upon the export industries, neither could it allow for the fact that, through economies of scale, the industries which initially were completely dependent upon the region for exports or for markets could develop interregional ties. For this reason it would be reasonable to consider Tiebout's version as the most plausible, but due to the predominance of exports within the exogenous sector, one would not wish to downgrade their importance in the short run.

The second conceptual difficulty arises from the fact that the multiplier as usually measured is a combination of two types of multipliers. The first is determined by the extent to which the final export sector contains or uses intermediate products manufactured locally. The second is strictly the Keynesian type dependent upon changes in local income flows which in this case is determined by the consumption habits of employees in the export, intermediate and service industries.

Although this may be considered more a conceptual refinement, rather than a conceptual difficulty, problems do arise for strictly the economic base concept implies the measurement of the latter multiplier whilst in practice what one is often measuring is only the former (viz. in the firm-by-firm approach).

Thirdly, and finally, like the sector theory the economic base theory is guilty of being based purely upon demand considerations, those of supply being ignored. The most serious consequence of this is the theory's neglect of productivity changes. Clearly any increase in demand for exports may lead to changes in productivity, not only in the exporting industries, but also in the intermediate and service sectors as well, and should there be a greater scope for such increases in one sector rather than another then the multiplier will obviously change accordingly. Another aspect of the supply problem is the failure of the theory to appreciate that investment is not only income-creating but is capacity-creating as well. Really the concept should embody some of the principles akin to the Harrod-Domar Growth Models which explicitly recognise these two aspects of investment.

Conclusions to regional multiplier analysis

The regional multiplier simply stated is a statistical relationship between the exogenous and endogenous sectors of a region. That it is valid both to divide the region's economy into these two sectors and to consider the exogenous sector as the 'basic' one has never really been disputed. But the economic base concept in particular goes further than this by maintaining that exports are the all-important constituent of the exogenous sector.

This is a highly suggestive, but simplifying, generalisation which lends to the theory the same characteristics as those of the sector theory, viz. that in aggregate it may be useful, but in any particular application it must be treated with great caution. The emphasis on exports rules out any long-run application and the considerable technical difficulties inherent in its application in the short run leads to the necessity for treating any conclusions that may emerge with great care, if not suspicion. But despite these objections the theory still has a remarkable vitality, and is indeed to be found in almost all regional forecasting studies. This is probably due to the facts that the base concept has not been seriously conceptually undermined by its critics, and as such remains valid, together with the knowledge that, despite its simplicity, it still remains the most sophisticated method that can yield results, given the general paucity of data available at the regional level. In short, as Meyer (35) has noted, there is no real empirically implementable alternative open to the regional analyst.

REMAINING ECONOMIC THEORIES

As stated at the beginning of this chapter, the remaining economic theories and techniques open to the regional analyst have minimum data requirements which far exceed the level of sophistication obtained at the regional or sub-regional level in the U.K. Yet although these theories cannot be applied to the problem at hand, it is nevertheless important for a full understanding of the subject to show briefly why this is so.

A thorough review of the data available, at the town level, together with a full discussion of why this level was adopted will have to await the empirical section of this thesis. At this stage it is sufficient to note that as far as the theories below are concerned, at least one of the following sets of data is vital to each of them; viz. data on inter-regional flows, inter-industry but intra-regional flows and inter-industry comparative cost levels. It is the absence of these statistics which rules out any approach based other than on the sector or economic base theories.

The most important exclusion on these grounds is perhaps that of the input-output technique, of which economic base theory is a highly simplified form. Although to a large extent this technique is merely a descriptive device, it nevertheless could be useful in forecasting industrial structures. Basically regional input-output analysis consists of a matrix of input-output coefficients relating not only to inter-industry flows but also to inter-regional flows as well. Consequently, on the basis of an assumed rate of growth in either final demand or the national economy, the relative expansion paths of each industry in any region could be calculated and, therefore, the change in any region's industrial structure forecasted. Clearly in its idealised form the input-output table's data requirements are far too sophisticated, and it is not surprising to find that most developments in this technique have been aimed at reducing the empirical detail required. Meyer (35) has illustrated the principal ways in which this has been achieved. Attempts have been made to aggregate the flows either by subsuming all inter-regional flows into one export/import sector or, conversely, aggregating over all industries and concentrating solely upon inter-regional relationships. But the former still requires data on inter-industry flows and the latter on inter-regional flows, neither of which are available. Perhaps the most promising simplification is that advanced by Leontief and Strout (30) who adopted gravity-type structural equations as proxies for inter-regional flows. Nevertheless the data requirements for this method still involve the output of the industries in each region, which again is data not available at the sub-regional level in the U.K.

The other important technique which has to be rejected is that of industrial complex analysis, which itself is an amalgam of input-output and comparative cost techniques. As already noted comparative cost analysis is used chiefly to identify the economic base of a region. Although only those costs which vary between regions need be considered, thus reducing considerably the demands made upon data, the method still relies upon detailed cost figures on those inputs which do vary between regions, principally transport, labour and a few key raw materials. However, the analysis is only partial and as a result it ignores the complexities arising out of economies of scale, factor price changes induced by inelastic supply functions and the economies resulting from agglomeration. Industrial complex analysis, as developed by Isard <u>et al</u> (28), overcomes some of these difficulties by attempting to define meaningful industrial complexes for cost comparison by using input-output matrices. Specifically the method calculates, on the assumption of mainland production functions, what would be the optimal petro-chemical complex for Puerto Rico, and then relaxes this assumption so as to allow for important non-identities between the mainland and Puerto Rico operations.

In order to justify the omission of both comparative cost and industrial complex analysis, it is necessary to beg the question of the nature of the empirical work to be conducted in this thesis. As will be shown later, either one could concentrate upon a particular region and forecast its industrial structure or one could conduct a cross-section study based upon a large number of small sub-regions and implicit in the discussion of input-output techniques was the adoption of the latter. Nevertheless it must be realised that in a limited manner national input-output coefficients could be used to build a model for any particular region and to this extent industrial complex analysis could be considered a real alternative. Yet if the aim is the more sophisticated one of analysing the general determinants of industrial structure (on the basis of a cross-section study) then clearly the data limitations must preclude the adoption of these techniques.

Consequently, as a result of data limitation, only the sector and the regional multiplier theories will be carried forward into the empirical section of this thesis.

Chapter Five

THE MODEL

The last three chapters reviewed those theories which at some stage have been used to explain differences in industrial structure or which by implication could have been so used. It is the aim of this chapter now to synthesise and develop the positive conclusions of the last three chapters into a model that can be used to derive the determinants of industrial structure. This model, which will be carried over into the empirical section of this thesis, must reflect that combination of aim, available data and theory which is capable of yielding the most meaningful results. Nevertheless, although the choice of this model in practice involves the simultaneous interaction between these three, for exposition purposes it will be derived in two separate stages. First, data considerations will be ignored and a general model developed consisting solely of those hypotheses which stem logically from those theories outlined in the previous chapters. Secondly, in the light of the data available, this model will be adjusted and adapted so as to be consistent with the data available to this particular study and it is only this latter model which need reflect the above combination of aim, data and theory.

AIM OF STUDY

However, before either of these stages can be tackled it is necessary to specify the aim of this research more precisely than has been hitherto attempted (see the Introduction). The aim arose out of a general awareness of the regional economists' inability to predict the changes in a region's industrial structure, and a general desire to improve upon this situation forms the starting point to this study. Conceptually one could achieve this either by restricting one's aim to a particular region and to analyse the forces which have dictated its development in the past and then, in the light of this and other factors peculiar to the region, to forecast the likely change in industrial structure. But the disadvantage of this case-study approach is that the conclusions are likely to be of little significance outside that particular region. For this reason the alternative aim of discovering what factors have influenced the development of a number of regional industrial structures will be adopted here. In this way the general factors at work may be elucidated which, if it is so desired, could be applied to any individual region. The disadvantage of this method is that it is unlikely to be able to predict the industrial structure of any one region as accurately as the first, but this cost in accuracy is outweighed by the benefit of greater generality and, consequently, overall usefulness.

However, until what is meant by 'industrial structure' is defined more precisely, the above aim will still be too wide for purposes of formulating hypotheses. It would be consistent with the above paragraph to consider either industrial structure in a disaggregated form based on SIC orders or in an aggregated form based upon the primary, secondary and tertiary sectors. Without any prior research to fall back upon one could reasonably postulate that the forces at work would first express themselves in the division between the various sectors, so that any finer division between the industrial orders could be interpretable only in terms of the initial division between these broader sectors. For this reason alone one would be inclined to adopt an industrial structure defined in an aggregate manner since it is obviously wise to study the simpler forces at work first. This point is further reinforced by the fact that, as shown in Chapter 4, the theories based upon economic considerations are likely to be operative only for the broader division in industrial structures.

Consequently, for this research, industrial structure will be confined to the division between the primary, secondary and tertiary sectors. This then serves to raise the final point of which of these sectors the hypotheses below should attempt to analyse. The primary sector can be discounted immediately as being of little significance. Of the remaining two, the tertiary sector would appear to be more susceptible to analysis since the theories outlined in Chapters 3 and 4 were, implicitly at least, more applicable to a study of that sector. This is clearly true for the sector theory, whilst the other economic approach, the economic base theory, treated the exporting sector (which is chiefly interpretable in terms of the secondary sector) as exogenous. Furthermore, central place theory was based upon the functions performed by regional economic units and it is really only in the tertiary sector that these are free to vary.

Hence, for this study, the aim will be to determine those forces which influence the size of the tertiary sector in regional economic units.

GENERAL HYPOTHESES

Given the above aim of this research, it is now possible to consider what contribution can be made by each of the theories outlined in the previous chapters. The theories forming the essential background to the model are diverse and, as a result, the various formulations derived from them can be thought of as being alternative ways of looking at the overall problem. The main division in the above theories lies between those based upon location theory and those stemming from macro-economic theory, for the former would attempt to explain the industrial structure of a regional economic unit in terms of its relationship with other such units and its place in an hierarchy, whilst the latter would study the unit in isolation within a strictly economic framework. This division has been carried into the selection of the independent variables and attention will be directed to both in turn, taking those hypotheses stemming from location theory first.

It will be recalled from Chapter 3 that location theory formed the theoretical base to central place theory, the main conclusion of which was that there would be a natural ordering of regional economic units in the form of an hierarchy. Lösch (32) derived this notion of a natural ordering from his concept of an ideal market and Christaller (7) showed how, as a result of economies of scale, the position in the hierarchy would be associated with the number of central goods produced in that place. In other words central place theory can be reduced to the hypothesis that

$$C_i = F[H_i] \qquad (5.1)$$

where C_i = the number of central goods produced by regional economic unit i

H_i = the position of regional economic unit i in the hierarchy.

As such the hypothesis has little apparent connection with the dependent variable, i.e. with the size of the tertiary sector. The theory merely states that the higher the grouping of an economic unit the greater will be the number of central goods produced there, but it shows no connection between a unit's grouping and its industrial structure. Consequently the theory has to be extended by the assumption of a real relationship between industrial structure and the number of central goods produced by a regional economic unit. If for the moment the definition of a 'regional economic unit' is begged and the town is taken as such a unit, then the assumption can be defended on the grounds that the tertiary sector of a town also serves the area around that town and ceteris paribus the larger the town the larger the area it serves and the more specialised become its services. In other words the larger the town the greater will be its tertiary sector due to the increased range of goods produced there i.e.

$$S_i = F\left[C_i\right] \qquad (5.2)$$

where S_i = the size of the tertiary sector of regional economic unit i.

The hypothesis in (5.2) will be valid to the extent that an aggregated definition of industrial structure is adopted, though it would break down in the case of the disaggregated form (i.e. a definition based upon individual industries). It is also an acceptable assumption so long as the town is taken as the basic unit, but a full discussion of this particular point will be left until later. So, accepting the aggregated form of industrial structure and leaving aside for the moment the question of which unit to adopt, hypotheses (5.1) and 5.2) may be combined to yield the third hypothesis that

$$S_i = F\left[H_i\right] \qquad (5.3)$$

This notion that the size of the tertiary sector of a regional economic unit is a function of its place in the hierarchy serves as the basic starting point for the ensuing discussions.

However, this hypothesis will be of little practical use unless either an hierarchy can be identified absolutely or some independent measure of an hierarchy can be found. In Chapter 3 it was shown that attempts to isolate an hierarchy in absolute terms were unsuccessful. However, studies which adopted population and function as independent criteria for an hierarchy were partially successful since both these methods found the presence of relatively few classes of towns each with a relatively large functional gradation. On the basis of these results it would appear possible to further develop hypothesis (5.3) through respecifying the hierarchy in terms either of the population of a regional economic unit or the function performed by it.

It will be remembered that the rationale for taking population as a criterion for the hierarchy lay in the fact that, as shown in Chapter 3, central place theory identifies groups of regional economic units according to the number of central goods produced there and states that each group is characterised by a certain minimum population. The latter point follows from the fact that the basic cause for grouping lies in economies of scale in the production of central goods and that a certain minimum population is required before these can materialise. As a result it should be possible to identify a regional economic unit's position in an hierarchy from the population associated with it. Given this fact, then it is possible to develop hypothesis (5.3) into the following

$$S_{ti} = F\left[P_{ti}\right] \quad (5.4)$$

where S_{ti} = the size of the tertiary sector of regional economic unit i at time t

P_{ti} = the population associated with regional economic unit i.

This hypothesis, that the size of the tertiary sector is a function of the population of the regional economic unit, is the first hypothesis to be incorporated in the general model.

To consider now the alternative criterion, that of function, in Chapter 3 Ullman (60) was shown to be one of the main protagonists of this idea in which he identified Christaller's (7) central places as market hamlets, township centres, county seats, district cities, small state capitals, provincial head cities and regional capitals. Consequently the industrial structure of a regional economic unit, or at least the size of its tertiary sector, could be identified by deciding which group the unit fell into.

The difficulty now presents itself of how to incorporate this into a testable hypothesis. One method would be to use dummy variables to identify the category which a town fell into. This would then lead to the following hypothesis in which

$$S_{ti} = F\left[X_a \ldots\ldots X_z\right] \quad (5.5)$$

where X_a = dummy variables representing a particular group in the hierarchy.

Thus dummy variable X_a could refer to a regional centre and it could be assigned a value of zero or unity depending upon whether the particular unit was, or was not, such a centre.

However, attempts to justify the claim for a distinct hierarchy, taking the town as the basic unit, occupied much of Chapter 3, the general conclusion there being that the claim was largely invalid. It will be recalled from this chapter that two approaches were adopted in trying to identify hierarchies in towns. The first analysed the relationship between the town and the rest of the system (viz. Carruthers (6) based his study upon bus routes) whilst the second tried to study a town in isolation through the adoption of a battery of statistical indicators (e.g. Smailes (51)). Unfortunately, although each method managed to generate its own hierarchy, when these were compared they were found not to correspond to each other and this was held not to be surprising in view of the weaknesses associated both with their criteria and assumptions. The consequences of this lack of correspondence between the hierarchies will be investigated in more detail later on, but all that need be noted here is that as a result it becomes impossible to identify the groupings in an hierarchy other than subjectively, and hence the claim to a clear and distinct hierarchy is invalid. This conclusion is important here, for it means that any approach based upon dummy variables has to be rejected, for a pre-requisite to such an approach is an ability to identify the various groupings in reasonably objective terms. This is quite apart from the conclusion to Chapter 3 that others have attempted this method and singularly failed for well understood reasons.

The fact that there is no evidence for a single unified hierarchy suggests that the strict interpretation of central place theory, in which population and function were taken as alternative methods of identifying the same hierarchy, should be rejected. But this does not imply that the findings of this chapter should be dismissed, for there is clearly too much evidence of a systematic variation between industrial structure and function for that to be permissible. An alternative way of looking at the problem would be not to postulate a connection between population and function, but rather to consider the functions performed by a regional economic unit as reinforcing or reducing the influence of population in determining the size of the tertiary sector. Hypothesis (5.4) states that the size of the tertiary sector of a regional economic unit depends upon its population, but Chapter 3 also shows that the function performed by the unit is important. Thus for two units of equal size the influence of function would be positive in the case of one being a regional centre and negative in the case of the other being a manufacturing town placed near to a conurbation. In the former case the tertiary sector would be larger than expected for the population alone, whilst in the latter it would be smaller. In other words hypothesis (5.4) now becomes

$$S_{ti} = F\left[P_{ti}, R_{ti}\right] \qquad (5.6)$$

where R = the function (rôle) of regional economic unit i.

It should be noted here that hypothesis (5.6) extends the concept of central place theory and at the same time manages to avoid one of its main pitfalls. As already argued, central place theory would consider both population and function as identifying the hierarchy in absolute terms, but this served only to raise the difficulty that, although it is possible to identify an hierarchy by both methods, the two fail to correspond to each other. Yet this difficulty stemmed solely from the attempt to identify in absolute terms the group a unit fell into. In fact hypothesis (5.6) shows that it is sufficient only to consider whether, on account of the function performed, a unit will be in a relatively higher or lower group than would be expected from its population alone. As a result hypothesis (5.6) not only extends the concept of central place theory but also avoids one of its major weaknesses. This, then, will be the second hypothesis to be incorporated into the general model.

The above concludes a discussion of the contribution made by location theory and its descendants and attention must now be directed to a development of the regional multiplier and sector theories as contained in chapter 4. It will be recalled that these two theories approached the problem through analysing the regional economic unit in isolation rather than, as in Chapter 3, as part of a more general system.

To consider first the regional multiplier theory, this would postulate that changes in one sector would be transmitted to another in the usual Keynesian manner. Conceptually this method could be used to analyse any number of sectors and a complex model could be devised upon the interaction between a chosen number of them. However, as shown in Chapter 4, it is a quite general characteristic of any regional unit that the data is insufficient for such a procedure. This has the effect of considerably weakening the application of this concept since it has to be simplified into the economic base concept to make it compatible with the data available. In addition to the weakness associated with the economic base concept (outlined in Chapter 4), as will be seen later, the method poses certain difficulties in specifying the relationships even at this level of simplification. At this stage, it should be recalled that, as Meyer (35) has already noted (Chapter 4), there is no real alternative to this method and that the only course left to the analyst is to bring into the open these weaknesses and to interpret the conclusions always with these in mind.

In its most general form, the economic base theory traces the inter-sector income flows between the exogenous and endogenous sectors according to two hypotheses:

$$Y = F\,k\left[X_i \ldots\ldots X_n\right] \qquad (5.7)$$

$$\text{and } Z = F\,Y \qquad (5.8)$$

where Y = the total income of the region

X_i = the output of the exogenous industry i

Z = the output of the endogenous sector

k = a constant

Thus the output of the exogenous sector determines a region's income and hypothesis (5.8) shows that this level of income is sufficient to support a given output of the endogenous sector. Thus combining hypotheses (5.7) and (5.8) gives the basic premise that the output of the endogenous sector is a function of the exogenous sector i.e.

$$Z = F\,k\left[X_i \ldots\ldots X_n\right] \qquad (5.9)$$

However economic base theory went further than this by postulating that the output of the 'exporting' sector of a region is the principal component of the exogenous sector's output. Thus hypothesis (5.9) can be further refined into

$$Z = F\, k_i \left[E_i \ldots\ldots E_n\right] \qquad (5.10)$$

where k_i = constant

E_i = the output of exporting industry i.

Nevertheless, in order to adapt this general conclusion of the economic base theory to the problem at hand, it is necessary to relate hypothesis (5.10) to the tertiary sector. This can be done through the legitimate assumption that the endogenous sector itself is principally identifiable in terms of the tertiary sector. However this only serves to raise the difficulty that hypothesis (5.10) relates to the whole of the exporting sector, part of which is to be found within the tertiary sector itself (i.e. within the dependent variable). Since this is illegitimate the exporting sector will have to be confined to those industries located within the secondary sector. Any resultant inaccuracy involved in excluding those exporting industries in the tertiary sector can largely be overcome by interpreting the results from this hypothesis together with those of hypothesis (5.6) in which the role of the tertiary sector as an exporting sector is explored. So, allowing for these adjustments, the third hypothesis to be adopted becomes

$$S_{ti} = F\left[E_{mi} \ldots\ldots E_{mn}\right]_{ti} \qquad (5.11)$$

where E_{mi} = size of exporting industries in the secondary sector.

Before attention can be turned to the sector theory it is necessary to fully understand the differences in working between this theory and economic base theory. In the latter theory, the causal relationship runs from the exporting sector, through the level of income and the regional multiplier, to the size of the tertiary sector, or in schema form

exporting sector $\rightarrow$ level of income $\rightarrow$ Regional multiplier $\rightarrow$ size of the tertiary sector

The sector theory, on the other hand, goes no further back in the process than the level of income in a region and traces the relationship between this and the size of the tertiary sector, not in terms of the regional multiplier, but rather in terms of differential sectoral income elasticities and productivity changes. Again in schema form,

level of income $\rightarrow$ differential sectoral income elasticities and productivity changes $\rightarrow$ size of the tertiary sector

Thus the sector theory would state that a rise in per capita income would lead to an expansion of the tertiary sector relative to the secondary sector due to the presence of a higher income elasticity in the tertiary sector, together with fewer opportunities to improve that sector's productivity. Consequently for two similar regional economic units, one with a higher per capita income than the other, <u>ceteris paribus</u> there will be a difference in the size of their tertiary sectors. Hence the fourth basic hypothesis to be incorporated in the general model is

$$S_{ti} = F\left[Y_{ti}\right] \tag{5.12}$$

where Y_{ti} = the level of per capita income in the regional economic unit.

<u>Form of the general model</u> - it is now possible to bring together the above four hypotheses to be tested into a single general model viz.

$$S_{ti} = F\left[P_{ti},\ R_{ti},\ (E_{mi}\ldots\ldots E_{mn})_{ti},\ Y_{ti}\right] \qquad (5.13)$$

In other words the size of the tertiary sector of a regional economic unit will be a function of its population, the rôle or function it performs, the size of its exporting sector and the level of per capita income within the unit.

The question now raised is whether this will be in a linear or non-linear form. It should be pointed out that it will be the task of the empirical section to discover exactly what form the relationship will take and that the main concern here is simply to postulate that model which theory would lead one to test. To the extent that the theory behind each variable has little to say upon the actual form of the relationship and that to adopt a non-linear form at this stage would perhaps obscure the essential arguments behind each variable, a linear form will be accepted. In which case equation (5.13) can be re-written as

$$S_t = a_1 + b_1P_t + c_1R_t + d_1(E_{mi}\ldots\ldots E_{mn})_t + g_1Y_t \qquad (5.14)$$

where a_1 b_1 c_1 d_1 and g_1 = constants.

This is not to deny the fact that the relationship might be other than linear, but at this stage it would be impossible to state exactly what non-linear form it would take. However, for the purposes of the empirical section, it should be noted here that if the size of the tertiary sector is expressed in percentage terms its value must be restricted to the range 0 to 100, in which case with a linear form it may well be possible that doubling the size of, say, population, would lead to a tertiary sector greater than 100%.

Up to this point the model has been developed in terms of the percentage size of the tertiary sector in one year, but it would appreciably add to the usefulness of the model if this could be adapted to explain changes in the size of the tertiary sector also.

One hypothesis which could be advanced is that the changes in the tertiary sector might be a result of changes in the independent variables within the same time period. Thus, for instance, a change in population over the period would cause a corresponding change in the size of the tertiary sector, i.e. equation (5.14) would then read as

$$\Delta S_t = a_2 + b_2\Delta P_t + c_2\Delta R_t + d_2\Delta(E_{mi} \ldots . E_{mn})_t + g_2\Delta Y_t \qquad (5.15)$$

Both equations (5.14) and (5.15) represent an equilibrium system in which the interaction between the dependent and independent variables is completed in the absence of any time-lags. In addition the first differences version of the model (equation 5.15) will also indicate the stability of the relationship, for the degree of fit obtained with this model will reflect the extent to which the relationship itself has altered.

However, the changes in the size of the tertiary sector may be brought about by changes in the independent variables rather than by changes within the same time period. In other words the system may be in a state of disequilibrium. Such a position may arise for two reasons, the first being a perfectly general one, whilst the second is peculiar to the theories associated with equation (5.14). To take the general point first it is clear that if the tertiary sector takes time to adjust to a change in one of the independent variables, i.e. if it is relatively inflexible, then at a given point of time the size of the sector actually measured will not be its 'true' size given the values of the other variables. In such a situation the model should be specified in terms of a movement towards the equilibrium position rather than in terms of that position itself. The second reason is related to this inflexibility and stems from the controversy over whether the economic base and sector theories are short or long-run influences. This controversy has already been surveyed in Chapter 4, and it is sufficient to note here that the sector theory, through its emphasis on income elasticities and induced productivity changes, is more likely to work in the long term, whilst either viewpoint can be taken over the economic base theory. The point here is that should either of these theories work in the long run then the size of the tertiary sector measured at any point of time will be responding to past changes in these variables, and this is the same as saying there will be a disequilibrium position.

The presence of a disequilibrium position means that it is not the absolute level of the tertiary sector that is a function of the absolute level of the independent variables, but rather the change in the tertiary sector over a period. In other words the true relationship is not that in equation (5.15) but rather

$$\Delta S_t = a_3 + b_3P_t + c_3R_t + d_3\left[E_{mi}\cdots\cdots E_{mn}\right]_t + g_3Y_t \qquad (5.16)$$

Since it is not possible a priori to state whether the system will be in equilibrium or not, it will be necessary to retain both the equilibrium and disequilibrium versions of the general model.

THE DATA

The general model derived was done so without any regard to the statistics available and, as a result, it has to be considerably modified in any particular instance to allow for data incompatibilities. This modification will be the task of the next section, but first the statistics available at the regional and sub-regional level for each of the variables in the general model must be reviewed.

1) Dependent variable - since the general model is mainly concerned with inter-sector flows, the size of the tertiary sector should be measured either in terms of income generated or output. But for the former there are simply no statistics available whilst for the latter the only output figures which exist are those relating to the retail trade [Census of Distribution 1961]. However statistics are available on the employment in each industry in each 'town' either from the Ministry of Labour returns (unpublished) or from the Census of Population 1951 and 1961.

2) Population - detailed statistics based upon the civil parish or ward are available in the Census of Population 1951, 1961 and 1966.

3) Function performed by the regional economic unit - apart from its industrial structure there are no statistics available upon the function performed by a regional economic unit, though as will be seen from the next section this is largely a result of the difficulty involved in defining what is meant by 'function'.

4) Exporting sector - conceptually the 'exporting' sector would be identifiable in terms of a regional input/output table i.e. one which showed both inter-industry and inter-regional flows, but less idealistically the sector could also be identified in terms of inter-regional flow together with a knowledge of a region's total output [i.e. the exporting sector would be identified by the percentage of the total output flowing

out of the region]. But inter-industry, inter-regional flows and total output statistics are not available and again resort will have to be made to labour statistics (see next section).

5) Income - statistics on total income disaggregated by source (viz. employment, profit and investment) are set out for most counties for certain years in the "Report of the Commissioners to the Inland Revenue", but below the county level reliance must be placed upon proxy-variables. [Statistics on the size of customers' deposits are published each year in the Trustee Savings Bank yearbook, but only for the 80 towns in which they have branches.] A discussion of the availability of these proxy variables, together with an assessment of their viability, will be fully developed in the next section.

One final point which is of considerable importance to the next section is the extent to which statistics differ between the regional and sub-regional level. The statistics so far outlined relate to the sub-regional level i.e. the town level and clearly these are also available through aggregation at the regional level. It might be legitimate to anticipate a considerable increase in the statistics available at the regional level, but this fails to materialise in practice. Apart from the statistics already mentioned, the major additions consist of activity rates by age and sex for the past 12 years, statistics on the sales and output in 1954 and 1958 for the secondary sector as a whole and details of total expenditure and ownership of consumer durables in the family expenditure surveys.

SPECIFICATION OF THE PARTICULAR MODEL

The need to adapt the general model to the particular circumstances of this study has already been noted and as a first step in this process the data available at the regional and sub-regional level was outlined in the above section. But before any detailed adjustments can be made to this model it is necessary to decide upon which method of analysis is to be adopted.

Given the above availability of statistics either factor or regression analysis would be suitable, the actual choice being based upon their respective compatibility with the overall aim of this thesis. Although one would never wish to reject irrevocably any particular method, on these grounds one may feel justified in casting the analysis initially within a regression framework. The aim of this study is to discover what factors will explain variations in the size of the tertiary sector and as such regression analysis is clearly relevant. On the other hand factor analysis, as was shown in Chapter 3, is a purely descriptive device. This is so because in this method a large number of characteristics relating to the unit under review are assembled and the analysis is then used to generate separate and independent components which will explain the greatest amount of variation in these characteristics. However the components so generated lack any behavioural content which restricts the analysis to being no more than a descriptive tool.

If the method to be selected is that of regression analysis then before the model can be specified it must be decided whether the study is to be based upon a cross-section or a time-series analysis. Since this is chiefly determinable in the light of the available data, which in turn is affected by the choice of the basic unit for analysis, this latter point must be decided first. From a review of the data in the last section it is clear that there are effectively only two levels of analysis viz. the regional (based on the standard regions) or the town. At the regional level there is a greater variety of data but unfortunately the choice of the region as the basic unit would make little sense in terms of the findings of Chapter 3. Both the population and location variables stem from the idea of an hierarchy of towns and whilst there is nothing absolute about selecting the town itself, it would clearly be illegitimate to think in terms of an hierarchy of standard regions.

To return to the question of a time-series or a cross-section study, the former has already been ruled out through the rejection of the case-study approach, but a choice still has to be made between a cross-section combined with a time-series approach or a pure cross-section study alone. However the selection of the town as the basic unit of analysis has made the former approach impracticable for the following reasons. The data on the dependent variable is available over a period of years from the Ministry of Labour (see above) which bases its statistics on the returns from the local employment exchanges, but the alternative source, the Census of Population, presents the data only for the two years 1951 and 1961. As will be seen below, the bulk of the data for the independent variables is taken from this census, the statistics of which are based upon the local administrative area. If the data for the dependent variables is taken from the Census of Population then there will be an automatic correspondence between the area covered by both the dependent and independent variables, whereas if the data for the dependent variables is taken from the Ministry of Labour source this will not be so. Consequently the practicability of adopting a cross-section-cum-time-series approach turns upon the inaccuracy caused by having the dependent and independent variables based upon different geographical areas. The Ministry of Housing and Local Government publish a map (25) which superimposes the employment exchange areas upon the local administrative areas and it is clear from this map that the areas fail to correspond. Furthermore it is impossible to adjust one set of data to the other since neither is presented in a more disaggregated form.

As a result of this difficulty it is possible only to conduct a cross-section study though, since some data for the dependent variable is given for both 1951 and 1961, the effect of time is not entirely excluded.

Accepting a method based upon a cross-section regression analysis, taking the town as the basic unit of analysis, then, in the light of the data available at this level, the general model will have to be re-specified in the manner spelt out below.

Dependent Variable

As already indicated the size of the tertiary sector should be measured in terms of output but unfortunately resort has to be made to employment figures. Further, from the discussion in the above paragraphs, it follows that the only period over which adjustments for a possible disequilibrium position can be made has to be restricted to that of 1951 to 1961. Consequently, the dependent variable can be defined as either the percentage of the total labour force of a town employed in the tertiary sector or the change in the percentage over the period 1951 to 1961.

Independent Variables

1) Population. It will be recalled from the first section that the first hypothesis to be considered is that the size of the tertiary sector is a function of population or, using the notation previously adopted,

$$S_{ti} = F\left[P_{ti}\right] \qquad (5.4)$$

The basic unit of analysis is the 'town' which, from an economic point of view, usually differs from the unit identified in the Census of Population [since the latter is based simply upon the area delineated by the local authority boundary]. Since the local authority area is invariably the smaller unit the required adjustment consists of bringing into account all the population units lying outwith the local authority area which could still be considered part of an homogeneous economic unit. Leaving aside the question of the definition of such a unit until the empirical section, the only difficulty that this adjustment raises stems from the illegitimacy of simply aggregating the respective sub-population units. This cannot be done since the distance between these units and the town centre acts so as to reduce the influence of the former.

The question of the influence of distance is important for it is raised also by the next independent variable. The only theoretical formulation which can be used to estimate the extent to which the influence of a population unit diminishes with distance is the gravity model, which was originally developed independently by Stewart (53) and Zipf (66). In its most general form the model states that the interaction (I) between two economic units of mass M_1 and M_2 is a function directly of their mass and inversely of the friction between them, (F), raised to some power n, i.e.

$$I = k \frac{M_1 M_2}{F^n} \tag{5.17}$$

where n = constant.

The actual variables used to represent the interaction, mass and friction will vary with the nature of the problem, each different specification leading to a different value for n. It follows from equation (5.17) that the influence I_p that a sub-population P_s will have upon a town of population P will be

$$I_p = \frac{k_p P_s}{d^u} \tag{5.18}$$

where d = the distance between town centre and sub-population unit s

k_p and u = constants.

It follows from equation (5.18) that the population variable P_{ti} in equation (5.4) must be replaced by the town's total effective population i.e. that equation (5.4) must be modified into equation (5.19) viz.

$$S_{ti} = F\left[P + \sum_{j=1}^{n} \frac{P_{sj}}{d_j^u}\right]_{ti} \tag{5.19}$$

2) Function Variables. The second hypothesis developed from the general model was that the size of the tertiary sector would be influenced

by the function or role of the town, i.e. that

$$S_{ti} = F\left[R_{ti}\right]$$

In Chapter 3 two main difficulties with this approach were identified. First it was seen to be impossible to identify absolutely which hierarchy a town fell into. Secondly no satisfactory criterion of function was developed in the studies reviewed in Chapter 3. It will be recalled that some took a single criterion (e.g. Carruthers (6) took bus routes), whilst some adopted a battery of statistical indicators (e.g. Smailes (51)), but all fell short of defining a satisfactory criterion since all were open to criticisms on both their definition and content. It has already been shown how the first difficulty can be largely avoided by identifying the unit a town falls into in relative terms only, whilst to solve the second problem an independent measure of function is required.

One approach which offers some scope in solving this problem has already been touched upon in the development of the above hypothesis. This hypothesis is based upon the 'relative' approach and as such it must be taken along with the first hypothesis i.e. that

$$S_{ti} = F\left[P_{ti}, R_{ti}\right] \qquad (5.6)$$

the argument being that the size of the tertiary sector as given by the first hypothesis will be increased or decreased according to the influence of the second. Further, in section one of this chapter, it was pointed out that the function performed by a town depended largely upon the presence of other towns nearby. Thus a town which itself was not large enough to be a regional centre in terms of population alone may well find that this function is induced by the absence of larger towns nearby. Conversely the presence of a larger town would tend to depress the size of that town's tertiary sector. In other words the size of the tertiary sector will be a function of the presence of larger towns nearby; the greater these are, the smaller the tertiary sector will be, i.e. that

$$S = F[L] \tag{5.20}$$

where L = influence of the population of larger towns nearby.

Conversely if there are a large number of smaller towns situated nearby their combined influence will result in the tertiary sector being expanded, whilst their absence would tend to depress the size of the tertiary sector, i.e.

$$S = F[M] \tag{5.21}$$

where M = index of the influence of the smaller towns nearby.

To briefly summarise the argument so far, the perfectly general hypothesis that the size of the tertiary sector is influenced by the function performed by the town can be developed into the hypothesis that the size of that sector will be a decreasing function of the influence of larger towns and an increasing function of that of smaller towns, i.e. combining (5.20) and (5.21),

$$S = F[L, M] \tag{5.22}$$

As a result equation (5.22) goes some way towards meeting the second difficulty outlined above, for an index of the influence of the larger and smaller towns can be used as a criterion of function, thus avoiding the difficulties of definition and content associated with those indices based directly upon the functional attributes of the town. To derive an index of this influence is now a task commensurate with the data available, for it has already been shown that, on the basis of the gravity formula, the influence (F) that a population unit (P) exerts upon a point distance (d) away is

$$F = \frac{hP}{d^n} \qquad (5.23)$$

where h and n = constants.

So a convenient index (F_L) of the influence of larger towns on any particular town will be given by

$$F_L = F\left[h_L \sum_{j=1}^{n} \frac{P_{Lj}}{D_{Lj}^{v}}\right] \qquad (5.24)$$

where P_{Lj} = the population of the j^{th} town larger than the town in question

D_{Lj} = the distance of the j^{th} town larger than the town in question

h_L and v = constants.

Similarly a convenient index (F_m) of the influence of smaller towns upon any particular town would be

$$F_m = F\left[h_m \sum_{j=1}^{n} \frac{P_{mj}}{D_{mj}^{w}}\right] \qquad (5.25)$$

where P_{mj} = the population of the j^{th} town smaller than the town in question

D_{mj} = the distance of the j^{th} town smaller than the town in question

h_m and w = constants.

So combining equations (5.24) and (5.25) the perfectly general hypothesis originally stated can be modified to suit the particular data available into the following

$$S_{ti} = F \sum_{j=1}^{n} \left(\frac{P_{Lj}}{D_{Lj}}\right)^{v}_{ti}, \sum_{j=1}^{n} \left(\frac{P_{mj}}{D_{mj}}\right)^{w}_{ti} \qquad (5.26)$$

3) **Export Sector Variable.** Equation (5.11) expressed the third hypothesis contained in the general model, as

$$S_{ti} = F\left[E_{mi}\ldots\ldots E_{mn}\right]ti \qquad (5.11)$$

i.e. that the size of the tertiary sector would be a function of the output of those secondary industries which export outwith the region. In order to re-specify this in a form suitable to this particular study, the two problems of a) deciding upon a measurement of output and b) identifying the 'exporting' industries must be solved. The first is easily solved by default, for as already shown there is no measure other than that of the employment associated with each industry. The alternative methods of identifying those industries which do export out of a region were reviewed in Chapter 4, the conclusion there being that the method based upon location quotients adopted by Hildebrand and Mace (22) was the only practicable one. It should be recalled that this method requires the harsh assumption of a coincidence between local and national consumption patterns, but since there is no viable alternative the only way to deal with this problem is to bear this limitation in mind when interpreting the results.

The location quotient is given by the formula

$$LQ = {}^{R_i}/_{R} \Big/ {}^{N_i}/_{N} \qquad (5.27)$$

where R_i = number employed in industry i in a region

R = number employed in all industries in a region

N_i = number employed in industry i at the national level

N = number employed in all industries at the national level.

It follows from (5.27) that if the industry is larger, in terms of the percentage of the whole labour force employed by it, at the regional level than the national level, it will have a location quotient greater than unity. An industry with such a location quotient will be above average in size and, granted the assumption of a coincidence between local and national consumption patterns, it must be exporting out of the region. So although the basic hypothesis needs no further adjustment, it can be seen that, in the light of the data restrictions, the exporting industries will be identified through location quotients and their size will be measured in terms of their labour force.

4) Income Variables. In the review of the data available it was shown that there is no direct measure of a town's income and that proxy variables will have to be used. But since precisely which variables will be selected itself depends upon further research, it is not possible at this stage to take the argument relating to the fourth hypothesis any further than equation (5.12).

Conclusions to the particular model to be used

From the above analysis it can be seen that as a result of data deficiencies equation (5.6) of the general model has to be replaced by more complicated equations. Thus equations (5.19) and (5.26) replace the original equation (5.6) though, as just stated, equations (5.11) and (5.12) need not be altered at this stage.

Hence the first version of the equilibrium model, equation (5.14), must now read as

$$S_{ti} = a_4 + b_4 \left[P + \sum_{j=1}^{n} \frac{P_{sj}}{d_j^u} \right]_{ti} + c_4 \sum_{j=1}^{n} \left(\frac{P_{Lj}}{D_{Lj}^v} \right)_{ti} + d_4 \sum_{j=1}^{n} \left(\frac{P_{mj}}{D_{mj}^w} \right)_{ti} + e_4 \left[E_{mi} \ldots\ldots E_{mn} \right]_{ti} + g_4 Y_{ti} \quad (5.28)$$

Or, in simpler notation,

$$S_{ti} = a_4 + b_4 P_{ti} + c_4 L_{ti} + d_4 M_{ti} + e_4 E_{ti} + g_4 Y_{ti} \quad (5.29)$$

where P_{ti} = the combined population and neighbourhood variable

L_{ti} = the index of the larger town's influence

M_{ti} = the index of the smaller town's influence.

Using this notation, the second version of the equilibrium model, i.e. equation (5.15), may be written as

$$\Delta S_{ti} = a_5 + b_5 \Delta P_{ti} + c_5 \Delta L_{ti} + d_5 \Delta M_{ti} + e_5 \Delta E_t + d_5 \Delta Y_t \quad (5.30)$$

Finally, the disequilibrium form may be written as

$$\Delta S_t = a_6 + b_6 P_{ti} + c_6 L_{ti} + d_6 M_{ti} + e_6 E_{ti} + d_6 Y_{ti} \quad (5.31)$$

EMPIRICAL SECTION

Chapter Six

THE DATA

In the theoretical section the model was developed and it is now the task of the empirical section to test this model and to present the conclusions stemming from this research. Clearly the first step in this process involves the collection of the data for each of the variables associated with the model. For those studies which rely more or less exclusively upon published data, although there may be a host of detailed considerations, the form in which the data is found is usually sufficiently consistent as between each observation that the whole subject matter can be dealt with in an appendix. However such a treatment is not so convenient in this study for there are no statistics which can be applied directly to the variables, so that, for each of the variables, the data must be either adjusted or processed. As a result the collection of the data is not straightforward since such a step involves the adoption of several important assumptions, all of which have a significant bearing upon the interpretation of the results. Consequently a discussion of the issues raised by the data forms an essential prerequisite to the interpretation of the results, and this is something which cannot conveniently be attempted in an appendix.

The aim of this chapter is therefore to review briefly the manner in which the data was collected, leaving aside for an appendix the minu-
[illegible] t~~i~~ae of this process. In the light of this review it will be possible to show how the data, through the necessity for certain assumptions, has imposed a particular form upon the model and to assess the implication

that problems associated with the data have for the interpretation of the results. In so doing the chapter has been divided into the selection of the sample and the collection of the data for the variables.

THE SAMPLE

In the last chapter the 'town' was accepted as the basic unit of analysis, which raises at this stage the three difficulties of how to define the 'town', the population and the sample.

The 'town' has been taken as that population unit which is associated with either county or municipal boroughs. A definition based upon local administrative areas was adopted since the bulk of the data used in this research was taken from the Census of Population which presents its information in local authority form. The difficulties and biases associated with this choice will be considered alongside the selection of the variables below.

The population was taken as all those towns possessing a population in excess of 50,000. Although the choice of this figure could be held to be somewhat arbitrary, there were nevertheless some good reasons for choosing this particular level. First it was clear from the rationale of the model developed in the last chapter that towns with small populations had to be rejected, since for these it would be meaningless to separate out the influence of larger towns (due to the presence of so many of them). Further, towns with a significantly smaller population would then be so small as to leave them with few tertiary activities that the larger town could usurp. Both these considerations in effect excluded towns below the 15-20,000 population level. Granted this, then the fact that the Census of Distribution gives more detail for towns with a population in excess of 25,000 implied that the lower limit should be fixed at least at this level. Further, in the publication "A new town in Mid-Wales" (64) it was shown that, as far as retail distribution is concerned, it was not until a theoretical shopping population of 35-40,000 was reached that there was a significant increase in the range of services provided over and above those associated with a town of a population of about 20,000. Since above this level there now appears an extra range of services that may or may not be provided depending upon the influence of larger and smaller population centres, it would seem sensible to consider only those towns capable of providing this extra range of services. If these three points are accepted, then the selection of a 50,000 level becomes convenient since it allows comparison with the major work in this field [Moser and Scott (37)] which adopted this population level.

The total number of towns with a population in excess of 50,000 in Great Britain was 178, and from these a final sample of 69 towns [see Appendix D Table D 1] was selected upon the basis of the key decision to exclude all those towns whose industrial structure was likely to be unduly influenced by the presence of London or a conurbation. As was shown in Chapter 5 one of the variables in the final model attempts to measure the effect that a larger town has upon the size of a given town's tertiary sector. However it is clear that London and conurbation centres (e.g. Liverpool, Birmingham, etc.) will themselves provide such an exceptional range of services that the relationship between them and their surrounding towns will be completely different from the relationship which may be presumed to normally exist. Thus the influence of London upon Croydon, Manchester upon Salford or Birmingham upon Dudley is likely to be of a completely different order from the effect of, say, Northampton upon Rugby or Brighton upon Worthing. Since there has been no previous research conducted in this field it is obviously preferable to tackle the simple issues first, and for this reason towns within a conurbation (as officially defined) or within the influence of London were deleted. The latter's influence was presumed to extend over a radius of 50 miles.

One possible source of error arose from the fact that similar distortions sometimes applied to two towns, both with populations greater than 50,000, which were close to each other, yet still outside London or a conurbation. If one town was significantly larger than another, viz. Brighton and Hove, Bournemouth and Poole, then it was thought reasonable to suppose that the larger town would also provide the more sophisticated services for the smaller town. In such an instance the smaller town could be relegated to the 'neighbourhood' variable (see below). However where the two towns were of equal size, e.g. Middlesbrough and Stockton, such an assumption was invalid and the only course open was to reject the towns.

THE VARIABLES

The following discussion on the collection of the data for each of the variables specified in the final model can be conveniently divided into two parts. The first will deal with the precise definition of each variable and the difficulties involved in working with that definition, whilst the second will consider the problems associated with the use of the data itself.

Dependent Variable

Definition - from Chapter 5 the dependent variable was defined simply as the size of the tertiary sector and in order to specify this variable more precisely two questions must be answered. First to what level of disaggregation should the data be taken and, secondly, should employment or output be taken as an indicator of size? With regard to the data it was presented with the minimum list heading as the smallest unit, but for the purposes of this thesis it was felt necessary only to disaggregate to the level of industrial orders. This was for the same reason that the towns under the influence of London and the conurbations were excluded, namely a desire to limit the analysis initially to a study of the major forces at work. In this respect it is obviously more important to study the causes for inter-town differences in the transport and professional and scientific industries rather than, say, the differences between accountants and lawyers.

The question of whether output or employment figures should be taken has already been touched upon in Chapter 5. As already indicated output figures would be the most significant measure of industrial size since the independent variables operate initially upon the demand for an industry's output, and this can only be accurately reflected in terms of employment if labour productivity can be assumed constant. However it has already been noted that, apart from retail distribution, no output figures are available at the town level, and this accordingly introduces a bias into the results to the extent that the above assumption may not be valid. Fortunately, the significance of this bias may be partially estimated from a comparison of the results obtained with the retail distribution using both employment and output figures.

Data - for this variable there are two sources of employment figures in the U.K., viz. the Ministry of Labour returns which are compiled from details of national insurance cards, and the Census of Population; both of which present their data in the same form but for different geographical units. Any discussion of the marginal advantages of collecting this data through a sample method, as in the Census, or through a counting of heads, as in the Ministry of Labour returns, becomes irrelevant when the wide disparity in the administrative units is considered. The data for the Census is based upon local authorities' administrative areas, viz. county and municipal borough, urban and rural districts, whilst the Ministry of Labour data relates to the local employment exchange areas. The advantage in using the Census data lies in the fact that it ensures an automatic correspondence between the geographical areas covered by the dependent and the independent variables, since for the latter all the data is either collected from local authority sources or takes the local administrative area as the basic unit. On the other hand, the adoption of the Ministry of Labour data would require the detailed changing of the boundaries of those areas associated with the data on the independent variables. This would be a dubious procedure for the population figures since the smallest unit, the ward or civil parish, is large in relation to the changes required, whilst it would be impossible for the remaining data which is presented for administration areas only. Nevertheless, if there was an acceptable geographical correspondence between the two units, the loss of accuracy would be worth while since data on industrial structure could be obtained for any year and would not be merely confined to the census years 1951 and 1961. This would then allow the 1966 Census of Population data to be used, giving a time span since the war of 15 years (1951-1966) instead of 10 years (1951-1961). Unfortunately a complete lack of such correspondence was shown in the Ministry of Housing and Local Government Administrative Areas map (25) which superimposed local employment exchange areas upon local authority administrative areas. For this reason reliance had to be placed solely upon the Census data for the employment figures, for not to do so would introduce an unacceptable loss in accuracy.

One difficulty experienced with both sources of data is that for 1961 the industrial structure is based upon the 1958 standard industrial classification, whilst for 1951 it is presented using the 1948 classification. Fortunately the main differences between these two lie within the secondary sector, but certain modifications have been introduced into the tertiary sector. Thus the main orders in the tertiary sector remain unaltered, but detailed changes have been made in the classifications of some of its components. The net effect of these modifications is that the change in the tertiary sector as a whole can be accurately measured, but that changes in the individual industrial orders may partly reflect changes in classification.

The individual components of the tertiary sector, as given in the Census of Population, are set out in Appendix A Table A2.

The data on retail output came from the Census of Distribution conducted in the same year as the Census of Population. In this Census detailed information is available upon the numbers employed, turnover, the particular type of establishment, its location within the town and the type of goods sold, but for our purposes it was sufficient only to take note of the aggregate turnover for each town. Since both the year for which the Census was conducted and the unit upon which it was based coincided with the Census of Population, no difficulty was experienced with this data. This data is set out in Appendix A Table A3.

Population and neighbourhood variables

It will be recalled that this was the first independent variable to be developed in Chapter 5, the essential hypothesis to be tested being -

$$S_{ti} = F\left[P + \sum_{j=1}^{n} \frac{P_{sj}}{d_j^u}\right]_{ti} \qquad (5.19)$$

Definition. The sub-variable P is straightforward, being the population lying within the county or municipal borough. As already stated in Chapter 5, the other sub-variable represents an attempt to ensure that the population measured is economically meaningful and is not merely confined within arbitrarily drawn local authority boundaries. Included in this sub-variable therefore is the population of those units which lie outside the boundary but within a chosen radius of the town's centre. For the purposes here it was assumed that any population within a radius of five miles formed part of the same economic unit, though for the reasons given in Chapter 5 this had to be weighted by its distance d_j from the centre.

The only difficulty associated with the definition of this variable arose from those situations in which the boundary of a large outlying population unit was within five miles of the town, but its centre was not. If the unit was compact then it was thought best to include the whole of it, whilst if it was of a rural nature only that population which from inspection appeared to be relevant was taken into account. This procedure was obviously somewhat arbitrary and could be a possible source of error, though the actual instances in which this difficulty arose were relatively small.

Data - the Census of Population 1961 gave the population figures for both the county or municipal borough and those wards or civil parishes that lay outwith the boundary, the latter being selected from an inspection of the quarter and one inch Administrative Area maps (25). The data for this variable is given in Appendix A Table A1.

Two sources of inaccuracy arose with this data. The first stemmed from the lack of data on the industrial structures of the wards and civil parishes. This was not relevant to those units which were not annexed to the town itself, since the aim here is to determine the extent to which the tertiary sector of that town was supported by the outlying population. But for those units which actually bordered upon the town it would obviously be desirable to include their employment, since it was only the arbitrariness of the local authority boundaries which excluded them in the first place. The second difficulty arose from the opposite situation to the first, viz. that some boroughs took within their boundaries a relatively large landward area (e.g. Swansea), the population of which would otherwise have been weighted by the distance from the town centre.

The cumulative effect of these two errors may lead to a bias in the results, a fact which will have to be borne in mind in their interpretation. Nevertheless, since the 'neighbourhood' effect is small in relation to the town's population, this bias may not be significant.

Function Variables

It will be recalled that the two separate hypotheses to be considered here are

$$S = F \sum_{j=1}^{n} \frac{P_{Lj}}{D_{Lj}^{v}} , \sum_{j=1}^{n} \frac{P_{mj}}{D_{mj}^{w}}$$

i.e. that the size of the tertiary sector is a function of the presence both of smaller and larger towns.

Definition - these variables give rise to two definitional problems. First what is meant by 'smaller' and 'larger' towns and secondly over what radius is their influence to be deemed to operate? The first is partly associated with the problem of whether to adopt a given percentage or an absolute population as a criterion for inclusion i.e. should a 'smaller' town be defined as one with a certain population or one with a certain percentage of the sample town's population. Both for the 'small' and the 'large' towns the absolute population was chosen, since it was felt that a town of a given population would be likely to exert an influence upon the sample town whether or not its population happened to be a certain percentage of that town's population. Granted this, then the first definitional problem resolves itself into a choice of this critical population level.

For the 'smaller' town variable the aim is not to measure all the population within a given area around the sample town, but rather to identify those units which not only could exert a significant influence upon that town, but which were also likely to vary as between the sample towns. For instance it would be unnecessary to include all those villages with, say, a population of 2,000 or more since each town would be characterised by a large number of such population units. Eventually, for the 'smaller' towns, a level of 10,000 was decided upon since towns of that size are relatively infrequent and also at this level they generally possess functions which could be assumed by a larger town. The situation is slightly altered in the case of the 'larger' town variable due to the impossibility of selecting any one critical level. This resulted from the fact that any given level would lead to the situation in which towns in the sample were both above and below this level, since the populations of the sample towns ranged from 50,000 to 500,000. This situation was resolved by accepting all those towns with a population greater than that of the towns under review.

The choice of over what radius the respective influences will operate must necessarily be somewhat arbitrary, but it is still clear that for the 'smaller town' variable, associated as it is with less sophisticated services, its influence will operate over a smaller radius than that of the 'larger town'. The influence of London has already been assumed to extend for 50 miles and it would be logical to assume a similar radius for the 'larger town' variable. Granted this, then it might be reasonable to assume the 'smaller town' influence as extending over the smaller radius of 30 miles (a generous limit was chosen deliberately since it would always have been easy to reduce this limit should subsequent research have shown this to be desirable).

There are two main difficulties in working with this definition. First, concerning the definition of the smaller towns, it is clear that this requires the assumption of a strict hierarchy of influence based upon population. If a town was smaller than the sample town but had a population of over 10,000 then it had to be included and this must lead to inaccuracies and biases where one town is regionally superior to, but smaller than, another. Such a town would normally be higher in the hierarchy than the larger town, but nevertheless under this system it still has to be subordinated to it, thus introducing a possible source of error (e.g. Cheltenham (pop. 72,154) and Gloucester (pop. 69,733)). The second difficulty arose from the need to reduce the radius under certain specific circumstances. These were the juxtaposition of two sample towns and the presence of complicating geographical factors. Examples of the first are

a) the smaller town variable - here it would be incorrect to include the influence of a smaller town when another sample town lay between the former town and the particular sample town under review, e.g. it would be incorrect to include the influence of Mexborough (pop.16,600) upon the sample town of Wakefield when another sample town, Barnsley, lay between them.

b) The large town variable - the radius had to be restricted where there were two equally sized population units one of which was significantly nearer to the sample town than the other, e.g. the influence of Manchester 31 miles from Barnsley was likely to be negligible in comparison to that of Leeds only 15 miles away.

The second cause of a restricted radius was that of complicating geographical factors, the main instance being that of the Pennines which effectively restricted the influence of Lancashire and Yorkshire towns to their respective county boundaries.

Data. The collection of the data was straightforward for both the smaller and the larger town variables since the population of the towns was given in the Census of Population 1961 and the distance of the towns was readily obtained from a road atlas. As with the population variable, the data for these two variables is set out in Appendix A Table A1.

Exporting sector variable

The relationship to be considered here is that

$$S_{ti} = F\left[E_{mi} \cdots\cdots E_{mn}\right]_{ti}$$

i.e. that the size of the tertiary sector will be a function of the size of the exporting industries in the secondary sector.

Definition - this has been considered in Chapter 5 in which an exporting industry was identified as being one with a location quotient greater than unity. The only definitional consideration raised here is that of which level of disaggregation to adopt. The broad category of industrial orders was selected as the basic unit, partly for the reasons outlined in the selection of this unit for the dependent variable and partly also to be consistent with that variable.

The main difficulty in working with this definition was also noted in Chapter 5, namely the necessity for an assumption of equivalence between local and national consumption patterns. At this stage there is little more to add to the remarks previously made.

Data - this has already been reviewed under the dependent variable heading and for obvious reasons of conformity the employment in each SIC order was taken from the Census of Population 1961, rather than from the unpublished Ministry of Labour returns. Granted a knowledge of the percentage distribution of labour force at the national level, it is easy to identify those industrial orders which have percentage distributions greater than this. Their values were then aggregated into a single value for each town. In Appendix A Table A4 the index so constructed is set out for each of the towns in the sample.

Income Variable

The relationship to be considered in this section is

$$S_{ti} = F\ Y_{ti}$$

or that the size of the tertiary sector is a function of the per capita income in the town.

Definition - since there is no data on income at the town level, proxy variables must be devised for a town's per capita income. In the past the search for such a variable has proceeded in several directions. One line of inquiry has been to look at the status and/or employment of the town's population, whilst a second has been to look at the ownership of assets which are associated with different income levels. Finally, a third method has centred upon the rateable value of the property within a town. Each of these methods will be looked at in turn.

In the first category one of the most popular methods has been to construct an index based upon the socio-economic classification given in the Census of Population, the rationale here being that the greater the percentage of employment in the higher status groups the greater will be the town's income. However, upon closer inspection, this method appears to have several deficiencies. First, although the categories are precise and narrow at the upper end, this is not so at the lower end of the spectrum, so that 55%-60% of the town's employment usually falls into the skilled, semi-skilled and unskilled categories. Hence it would not necessarily follow that a higher percentage of employment in, for instance, the first four categories would lead to a higher total income, since the latter is more likely to be determined by the percentage in the lower categories. Secondly, even if the index could be deemed to be sufficiently sensitive, a problem would still arise in interpreting the results since the income levels associated with each status group are not known. Thus it would not be possible to be certain that any difference in the income of a town, as a result of its distribution within the socio-economic spectrum, was sufficiently large vis-à-vis the others to influence the tertiary sector. In other words the socio-economic index may be related to something other than income, a point which is borne out by Moser and Scott (37) (see Chapter 3) who found that this index fell under a different component to those of the selected economic indicators. A final weakness of this approach is that a more sensitive indicator of the income generated in the lower groups would be given by the industry in which they worked rather than their particular socio-economic status. Thus more variation in income is likely to result from the predominance of a high wage earning industry per se, rather than the particular distribution of workers between the skilled, unskilled categories etc.

This last weakness of the socio-economic group would appear to suggest that a more sensitive measure of an individual's income would be to take the percentage distribution of employment in the town and, by weighting this with the actual earnings in that industry, to construct an index of earnings directly. This method would go some way to meeting the first criticism above, namely the impreciseness of the various categories into which the individuals would fall, since the index could be based upon the individual industrial areas. In addition the index would be more susceptible to interpretation since the income levels with which it is associated are known. In other words this index is at least partly based upon income data which the socio-economic index is not. Finally, since the method is based upon industrial structure, the third criticism is met, whilst the difference in the skilled, unskilled distribution would be partly manifested in the level of earnings applied to each industry.

It is for these reasons that an index based upon industrial rather than socio-economic structure has been preferred. However it should be noted that this index can only be based upon the primary and secondary sectors since the size of the tertiary sector itself will be the dependent variable. In addition, in order to avoid any spurious results in a regression analysis against the tertiary sector, the total of the individual orders times the wage earnings must be divided by the size of the primary and secondary sectors.

To consider the second category above, namely the assets which one might expect to be associated with different levels of income, the two to be included here are car ownership and television sets. Although strictly speaking it is not the number of cars but rather the expenditure on them that will vary with income, it may nevertheless be presumed that those towns with a higher per capita income will have a larger number of cars per population. With television sets the relationship may be expected to be even closer since the variation in types of sets is much smaller.

The third approach mentioned above is that based upon rateable values, and this finds its connection with per capita income through the workings of commercial law. The leading authority on this subject is edited by Mackay and Clyde (33). The actual determination proceeds upon rules laid down in case law which themselves stem from the statutory principle that:

> 'they must proceed from an estimation of the rent which a hypothetical tenant would pay to a hypothetical landlord in terms of the statutory tenancy in respect of the subjects for which the value is being determined'. Mackay and Clyde p.265.

Or, in economic terms, the rateable value must be based upon and must reflect the economic rent of the property. In the case of industrial and other commercial property the highest rent which an entrepreneur would pay would depend upon the income produced by the property, whilst for domestic dwellings it would be determined by the tenant's ability to pay. Hence on economic grounds it would be legitimate to pre-suppose a connection between income and rateable value.

In addition, the actual determination of rateable values must be comparable between towns since in practice rateable values are calculated upon one of the following principles,

a) the actual rent charged for the property

b) the rent charged upon comparable properties (Comparative Principle)

c) a 'standard' return based upon the cost of construction (Contractors Principle)

d) the profit earned from the property (Revenue Principle).

It follows that the same principles must be applied in each town, the actual method used depending not upon the town but rather upon the nature of the property. Furthermore, since there is a right of appeal to a central court, each of these principles has to be applied consistently. Both these considerations therefore lead to the conclusion that the relationship between rateable value and income must remain constant throughout the sample towns.

Rateable value therefore reflects the output of industrial and commercial property and, as such, one alternative form of specifying this relationship might be to consider the power consumption of each town. A priori the greater the industrial and economic activity the greater will be the power required by that town, so that one might expect there to be a relationship between fuel consumption and per capita income.

Data - for the index of earnings in the primary and secondary sectors, the industrial structure for each town is given in the Census of Population (see above) whilst the weekly earnings for each industrial order are given in the Monthly Digest. In this connection it is unfortunate that regional weekly earnings figures are not available since such a refinement would have been most valuable in this situation.

Details of car ownership for the year 1966 are to be found in the Census of Population 1966. The chief problem with these statistics is that they relate to 1966 and not 1961, so that an assumption about their relative movement through time is required. Details of the television licences in force for the sample towns was provided by the G.P.O. from unpublished statistics.

The data for rateable value is given for each county and municipal borough divided between domestic, industrial, shopping, office and other property in 'Rates and Rateable Values' (H.M.S.O.). The main difficulty with this data lies in the non-uniformity of its presentation, for prior to the year 1963-64, the data is given in much less detailed form. With regard to fuel consumption the only figures available were from the central electricity board's own records. Similar figures for either gas or coal were not available. The data for all the proxy variables for income are given in Appendix A Table A4.

Summary - from the above discussion it would appear that the proxy variables for income consist of an index of earnings, the numbers of either the cars per population or the television licences per population and, finally, rateable values and/or electricity consumption.

CONCLUSIONS

It was stated at the beginning of this chapter that since none of the data could be applied directly to the variables without the necessity of making important assumptions, the collection of the data could not be treated solely in an appendix. As a result, before the findings of the ensuing research can be interpreted, it is necessary to make explicit the assumptions adopted in the data collection, and, therefore, this chapter has attempted to set out those problems associated with this data. It has shown how these difficulties affect all the variables, and how especial attention must be paid to the composition of the sample, the function variables and the proxy variables for income.

Chapter Seven

THE RESULTS I

The model as postulated in equations (5.29), (5.30) and (5.31) of Chapter 5 will now be tested and the parameters estimated. However before this can be done preliminary investigations must be conducted to determine the best proxy variables for income and the correct weighting factors to be adopted for the population and geographical variables.

PRELIMINARY INVESTIGATIONS

To take the proxy variable for income first, it will be recalled from the last chapter that the possible sources lay in rateable value, car ownership, television licences, electricity consumption and an index of earnings per employee in the secondary sector. The data for income is assembled by the Inland Revenue for the county level but, unfortunately, it is impossible to compute statistics for television licences and electricity consumption at this level due to the incompatibility between their respective administrative units and that of the Inland Revenue. Hence these variables may only be evaluated through a regression against the size of the tertiary sector as the independent variable.

The results using the county income data as the dependent variable are set out in Appendix B, Table B1. The per capita rateable value was tried in total and also disaggregated across domestic property, shops, offices, other commercial property, industry, crown and other property. All the rateable values performed better against total income than against income from employment alone, giving for both an upward sloping relationship. For the former the best $\bar{R}^2$s were obtained for shops (0.12), offices (0.37) and other property (0.19), suggesting that an important distinction could be drawn between total non-industrial and total non-domestic rateable value, and indeed when this division was adopted the $\bar{R}^2$s were 0.02 and 0.63 respectively.

For car ownership it can be seen from Appendix B Table B1 that neither the number of households with one or two cars respectively, nor the number of cars per population exhibited any satisfactory relationship.

When the index of earnings is taken as the independent variable, the highest $\bar{R}^2$ was obtained against total income as the dependent variable (Appendix B Table B1), the linear relationship giving a marginally superior result to the log version (0.30 and 0.29 respectively). The sign of the coefficient was positive.

From Table B1 it can be seen, therefore, that non-domestic rateable value and an index of earnings are both significant proxy variables and that from a consideration of the strength of the relationship alone, the former would appear to be the most satisfactory. Scatter diagrams for both these variables reinforce this point (Appendix C Tables C1 and C2). The question nevertheless raises itself of whether either of these will be better variables when transferred to the town level. From Diagram C1 the relationship is more pronounced at lower levels of per capita non-domestic rateable values which suggests that, since towns typically have lower per capita values than counties, the relationship will hold there as well. Nevertheless, conceptually the index of earnings should be easier to translate to the town level since the procedure of multiplying the employment distribution by the national wage earnings is as likely to be valid at the town level as the county level, whilst there is no guarantee that a change in emphasis between the various methods of computing rateable values (see Chapter 6) would not introduce a disruptive element between these two levels.

As already noted the remaining two variables could not be tested at the county level, so that the only way these may be evaluated is to regress them against the size of the tertiary sector itself. However, when this is tried, as can be seen from Appendix B Table B2, no relationship is observed.

To turn to the second preliminary investigation, it will be recalled that the population and geographical variables were defined (using the notation of Chapter 5) as follows:

$$St = a_3 + b_3 \left[P + \sum_{j=1}^{n} \frac{P_{sj}}{d_j^u} \right] + d_3 \sum_{j=1}^{n} \frac{P_{Lj}}{D_{Lj}^v} + g_3 \sum_{j=1}^{n} \frac{P_{mj}}{D_{mj}^w}$$

The first step, therefore, is to identify which values of u, v and w give the best fit against the dependent variable. Values of 1.0, 1.5, 2.0, 2.5, 3.0, 3.5 and 4.0 were tried together with a value of u = 0 (corresponding to deleting the neighbourhood effect). The results obtained by using the values so generated as independent variables singly against the size of the tertiary sector as the dependent variable are set out in Appendix B Table B3. From this table it is clear that the values for u and v have no influence upon the overall relationship which either means that for the population and small town variables the gravity formula adds no sophistication or that it does but that this is masked by an overall lack of relationship. In either case the results show that the population of the towns as defined by their local authority boundaries are as accurate as any other measure, whilst the position is similar for the unweighted distance on the small town variable. An unweighted distance is also adopted for the large town variable, though this time Appendix B Table B3 shows clearly that a value for v of unity is the most satisfactory.

Conclusion - that as proxy variables for income only non-domestic rateable value and an index of earnings may be considered further. Of these two, rateable value gives the stronger relationship but is, nevertheless, less likely to apply unmodified at the town level. For the population and geographical variables, it would appear that a value of zero should be taken for u (population) whereas a value of unity for both v and w would appear to be the most satisfactory.

THE RESULTS

From the above preliminary investigations, taking both non-domestic rateable value (R_t) and the index of earnings (T_t) as proxy variables for income, the equilibrium model may be specified as either:

$$S_t = a_4 + b_4P_t + c_4L_t + d_4M_t + e_4E_t + g_4\ (R_t\ T_t) \quad .. \quad .. \qquad (7.1)$$

or $\Delta S_t = a_5 + b_5 \Delta P_t + c_5 \Delta L_t + d_5 \Delta M_t + e_5 \Delta E_t + g_5 \Delta(R_t T_t)$ ..(7.2)

Similarly the disequilibrium model would be

$$\Delta S_t = a_6 + b_6P_t + c_6L_t + d_6M_t + e_6E_t + g_6(R_t\ T_t) \quad .. \quad .. \qquad (7.3)$$

The choice of R_t or T_t as proxy variables for income will be made in the light of the ensuing results. Furthermore it should be noted that the population and neighbourhood variable (P_t) is simply the population of the town, whilst L_t and S_t in their full form are as follows (using the notation of Chapter 5):

$$L_t = \sum_{j=1}^{n} \frac{P_{Lj}}{D_{Lj}} \quad \text{and} \quad S_t = \sum_{j=1}^{n} \frac{P_{mj}}{D_{Lj}}$$

In this study the above equilibrium and disequilibrium models will be evaluated in two stages. First the independent variables will be run singly against the respective dependent variables to see whether there is any relationship by themselves. Secondly the variables will be added into the model successively and a multiregression analysis performed in order to (a) specify the full model, (b) discover whether any of the independent variables combine with each other so that they are more effective when taken together and (c) arrive at a decision about which variables should be rejected. This last point perhaps needs some amplification. If a variable singly and in multiregression shows no relationship its deletion is straightforward (unless the behaviour of the residuals suggests otherwise). However the variable by itself may well be significant yet when combined with the other variables may fail to be so due to collinearity between it and the other independent variables. In every regression analysis of this nature some collinearity may be expected (part only of which may be detected through an inspection of the correlation coefficients), but if the aim of the analysis is to specify the overall model, i.e. to derive the best fit for the model, then it is still legitimate to delete a variable which, when in a multiregression analysis, is found to be insignificant. The only provision would be that the results should be invariant with the order in which the variables are added.

Finally, as a matter of presentation, the results for each regression performed (in terms of $\overline{R}^2$) are given in Appendix B, whilst a discussion of the most important of these, together with an analysis of the relationships between the coefficients and standard errors, will be given in the text.

Equilibrium model - this model will be tested for both equations (7.1) and (7.2). For the former the first stage of analysis consisted of investigating the individual relationships between the dependent and independent variables and here the results can be best summarised in terms of the correlation coefficients obtained. (The individual regressions are given in Appendix B Table B4.) These are reproduced in Table (7.1) below, from which several interesting results are obtained.

TABLE (7.1)

Correlation Coefficients with the percentage size of the tertiary sector

VARIABLE	CORRELATION COEFFICIENT
Population	-0.06
Small Town	-0.02
Large Town	-0.47
Export Sector	-0.58
Index of Earnings	-0.65
Non-domestic rateable value	-0.05

It would appear that the population, small town and non-domestic rateable value variables are not related to the percentage size of the tertiary sector 1961, whilst, of the remaining three variables that are, only that of the large town carried the 'expected' sign. Thus, for this variable, a correlation coefficient of -0.47 illustrates the inhibiting effect of the presence of nearby large towns on the tertiary sector. However, for the export and index of earnings variables, although the coefficients all suggest the presence of a definite relationship, the sign is negative. Since this is at variance with traditional theory, this point must be returned to in detail later on.

In the light of these results, the specification of the model can be approached through the successive addition of these variables into a multiregression analysis and through observing the significance of the coefficients and the behaviour of the $\bar{R}^2$. This step resulted in equations (7.4) to (7.8) below.

If the index of earnings and the large town variables are taken together, equation (7.4) is obtained.

$$S_t = 9833 - \underset{(0.003)}{0.013}\, L_t - \underset{(2303)}{15718}\, T_t \qquad (7.4)$$

In equation (7.4) both the coefficients are significant at the 95% confidence level and the $\bar{R}^2$ is 0.51. With the addition of the third variable, the size of the export sector, the $\bar{R}^2$ increases to 0.62 and all the coefficients are significant. Thus:

$$S_t = 10529 - \underset{(0.003)}{0.013}\, L_t - \underset{(2228)}{11581}\, T_t - \underset{(0.06)}{0.27}\, E_t \qquad (7.5)$$

But when the population variable (equation (7.6)), the small town variable equation (7.7)), and the non-domestic rateable value variables (equation (7.8)) are taken into the model, the overall fit either remains the same or actually falls 1% (equation (7.8))

$$S_t = 11759 - \underset{(0.003)}{0.013}\, L_t - \underset{(2274)}{11310}\, T_t - \underset{(0.06)}{0.27}\, E_t - \underset{(170)}{113}\, P_t \qquad (7.6)$$

$$S_t = 11559 - \underset{(0.003)}{0.013L_t} - \underset{(2258)}{11480T_t} - \underset{(0.06)}{0.27E_t} - \underset{(168)}{110P_t} + \underset{(0.12)}{0.049M_t} \qquad (7.7)$$

$$S_t = 12007 - \underset{(0.003)}{0.013L_t} - \underset{(2278)}{11396T_t} - \underset{(0.06)}{0.27E_t} - \underset{(168)}{112P_t} + \underset{(0.12)}{0.075M_t} + \underset{(10.5)}{4.38R_t} \qquad (7.8)$$

From the relationship between the coefficients and the standard errors it can be seen that the large town, export and index of earnings variables always remain significant irrespective of the other variables, whilst the population, small town and non-domestic rateable variables are always insignificant and fail to improve upon the overall fit of the model. Finally the negative coefficient of the export and index of earnings variables is invariant with the variables included. It would appear therefore that the large town, index of earnings and export variables are significant by themselves but when they are taken together they have a greater combined explanatory power. On the other hand, the population, small town and non-domestic rateable value variables have no explanatory power by themselves, nor would their combination appear to provide any.

From these results it would appear that three problems arise. First certain variables are apparently insignificant which poses the question of whether they should be dropped from the model altogether. Secondly, the sign of the export and index of earnings coefficients is negative which is contrary to traditional economic theory. Thirdly, the overall variance explained (62%) is rather low suggesting that the cause of this should be further investigated.

The variables which from the above results should be rejected are the population, small town and non-domestic rateable value variables, but before a decision can be made on these it is necessary first to construct scatter diagrams to see whether there is a further possibility for any relationship. These diagrams are set out in Appendix C, Diagrams C3 and C4. It is clear from Diagrams C3 and C4 that the small town and non-domestic rateable values offer no scope for improvement. For the former this may be surprising but for the latter the difficulty of transferring to the town level has already been noted.

The scatter diagram for the population variable is more hopeful (Appendix C Diagram C5) for it is clear that an upward trend is destroyed by certain groups of extreme values. This suggests that if these were due to the influences of either other variables in the model or some common exogenous characteristic, then population could be considered as an explanatory variable. If the non-extreme values are considered it is clear that thirty-eight observations accorded with an upward semi-log relationship, leaving thirty-one observations which need to be explained by other factors.

From an analysis of the remaining towns it would appear that they fall into four categories:

(a) eight towns with a low tertiary sector which were associated with high large town ~~variables~~ val, plus five with a large tertiary sector and low large town values. These can be considered as falling into one group of thirteen towns, the extreme values of which are explicable in terms of the geographical variable.

(b) four towns which appeared to fit into the general relationship but which had rather low values for the tertiary sector.

(c) the holiday resorts, all of which had large tertiary sectors.

(d) three extreme towns, Chester, Scunthorpe and Sheffield, which had exceptionally high or low values for no apparent reason.

On the basis of these four categories it is possible to construct four samples i.e. (1) a sample of the thirty-eight towns without extreme values, (2) this sample plus those thirteen towns responding to the geographical variable i.e. a sample of 51, (3) these 51 towns plus four towns with rather low tertiary sectors i.e. a sample of 55. In practice samples 51 and 55 gave such similar results that sample 51 could be dropped. (4) the full sample of 69 towns, consisting of the 55 towns, plus the holiday resorts, plus the three towns for which there is no apparent explanation. The towns falling into the 38, 55 and 69 samples are set out in Appendix D Table D2.

When the above single and multiregression analysis was conducted with these three samples the results in Appendix B Table B5 were obtained. The main points to note are:

(1) The population (log) is only important for the 38 sample. For this, equation (7.9) was obtained

should
he logs

$$S_t = \underset{(158)}{948}\, P_t - 5497 \tag{7.9}$$

But for sample 55, the following equation resulted

$$S_t = 2014 + \underset{(236)}{279}\, P_t \tag{7.10}$$

and, for sample 69, equation (7.11) was

$$S_t = 5661 - \underset{(0.002)}{0.0008}\, P_t \tag{7.11}$$

It would seem that the presence of any viable relationship at all is very dependent upon the sample chosen. The marked difference between both the size and the sign of the coefficients (together with their relationship with the standard errors) in equations (7.9), (7.10), and (7.11) is rather disappointing.

(2) the large town variable is important for samples 55 and 69 only
(3) taking the sample most favourable to population (sample 38) resulted in an overall fit of 0.60 when all the variables were considered (Appendix B Table B5). It will be recalled that this was a smaller $\bar{R}^2$ than that obtained with the full sample (i.e. equation (7.8) has an $\bar{R}^2$ of 0.62).

From these results it is not clear precisely what conclusions should be drawn concerning the population variable. On the one hand it is possible to explain the behaviour of the extreme values which destroy the relationship in the full sample and when these are removed the population variable assumes considerable importance. Yet, on the other hand, the overall variance explained in the model is worse than that for the full sample. Since to adopt the 38 sample would restrict the generality of the conclusions and since it behaves no better than the full sample, it would appear sensible to accept the latter sample and therefore to reject the population variable. Nevertheless, the fact that these results can be obtained with the smaller sample suggests that the final conclusions should at least be cast in the context of the sample chosen.

One final point on the population variable which can be taken from these results is the conditions under which population might be expected to influence the percentage size of the tertiary sector. From the behaviour of the 38, 55 and 69 samples it would appear that population will be important in the absence of any exceptional characteristics in the town (e.g. it being a holiday resort, steel town, etc.) together with the absence of any extreme values in the large town variable (both high and low values). Although the presence or otherwise of these variables may present problems of definition, nevertheless they should give an indication of the kinds of situations in which the population variable may be important.

Thus from this analysis, together with the evidence of the scatter diagrams, it would appear that the model as given in equation (7.1) should contain only the large town, export and index of earnings variables. In which case the formulation given in equation (7.5) would appear to be the correct specification of the first version of the equilibrium model, i.e.

$$S_t = 10529 - \underset{(0.003)}{0.013}\, L_t - \underset{(2228)}{11581}\, T_t - \underset{(0.06)}{0.27}\, E_t \qquad (7.5)$$

To take the second difficulty posed by these results, namely the negative coefficients on the export and index of earnings variables, the main question to be decided is whether reliance may be placed upon these results. Here two possible lines of inquiry are suggested. First, do the variables measure the same phenomenon, for if they do then clearly the possibility of both results being erroneous must be reduced? Secondly, might the aggregative nature of the dependent variable result in an oversimplification of the forces at work, i.e. might several diverse relationships be present at the individual industrial level?

However, before these lines of inquiry are pursued, the obvious question of whether the particular specification of these variables imposes a negative coefficient upon the results must be met. In this particular instance the importance of this is heightened by the fact that both variables are based upon the industrial structure of towns. The main danger in this respect has already been met through dividing the variables by the size of the secondary sector. Obviously without this step a negative coefficient would be a definite possibility, since the greater the primary and secondary sectors (index of earnings) or the components of the secondary sector (export sector) ceteris paribus the smaller would be the tertiary sector. The question of whether, despite this step, there still can be a negative sign imposed upon these regressions has been explored in Appendix E, from which it is concluded that, under certain conditions, this might be possible. Nevertheless, all the evidence suggests that these conditions are not met in this analysis and that the only instance in which the conclusions to these variables must be treated with caution would be in the limiting case of no population change. Hence, although the conclusions to Appendix E must always be borne in mind, it is still legitimate and necessary to pursue this negative coefficient further.

To turn now to the first line of inquiry above, there are strong a priori arguments for presuming that the variables measure the same influence, since both were designed to reflect income either directly (index of earnings) or indirectly (export variable). If this presumption is correct then it could legitimately be expected that the two variables would be related. But, unfortunately, the results are ambiguous in this respect. If the export variable is taken as the dependent variable, then the index of earnings variable explains only 17% of its variance, i.e.

$$E_t = 2548 + \underset{(3994)}{15219}\, T_t \qquad (7.12)$$

Yet, on the other hand, the coefficient on the index of earnings variable is significant even at the 99% confidence level. Moreover, with a positive coefficient, the two variables show the type of relationship that would be expected if they were both reflecting the same factor.

One further point of evidence lies with the scatter diagrams of both variables against the size of the tertiary sector (Appendix C diagrams C7 and C8). If the variables are not related it would be expected that these diagrams would exhibit little resemblance. Yet both show a strong downward trend and although their extreme observations are different, there are relatively few of these in either case.

It may therefore be tentatively assumed that since the variables are significantly related and since the behaviour of their scatter diagrams are similar both variables are measuring the same influence, namely some aspect of per capita income.

The second line of inquiry suggested above involves considering various disaggregated versions of the dependent variable. However this point may be conveniently considered alongside the third difficulty posed by these results (the low overall variance explained), for this also raises the issue of disaggregation. If the overall variance explained is low this suggests that different forms of the dependent variable should be tried to see whether the true relationship is significantly different at the individual industry level.

This problem was pursued in two ways. First the independent variables singly and then in combination were tried against the various industrial orders comprising the tertiary sector (industrial orders 18 to 24) then, secondly, in the light of these results various combinations of these orders were adopted as the dependent variable. The results are set out in Appendix B Table B6 and B7.

To consider the export and index of earnings variables first, the results of both the single and multiregression analysis for the various industrial orders singly are set out in Appendix B Table B6. The most important results were for the public utilities, distribution, finance and public administration sectors and these are reproduced in Table 7.2 and 7.3. For, first, the index of earnings variable, column 4 of Table 7.2 shows that the best fit is still obtained with the total tertiary sector as the dependent variable, although as can be seen from an

Table 7.2

Specification of equations obtained with individual Industrial Orders - Index of Earnings Variable

Industry	Constant	Coefficient	$\bar{R}^2$
Public Utilities	254	-208 (169)	0.07
Distribution	2610	-3977 (754)	0.28
Finance	1715	-2977 (753)	0.17
Public Administration	1117	-2112 (936)	0.05
TOTAL TERTIARY SECTOR	10083	-18381 (2422)	0.45

inspection of the coefficients and standard errors in column 3, only the public utility sector was actually insignificant at the 95% confidence level. Moreover it can be seen that it is at this level that the relationship is the most significant. Precisely the same results were obtained with the export variable (Table 7.3) though this time the public utility category is significant. Thus Tables 7.2 and 7.3 seem to indicate that the total tertiary sector is not too aggregative a variable, and that it does not hide any fundamentally different relationships at the individual industry level. Hence Tables 7.2 and 7.3 may be taken as further evidence that some reliance can be placed upon the coefficients obtained with the export and index of earnings variables in equation (7.5).

Table 7.3

Specification of Equations obtained with Individual Industrial Orders - Export Sector Variable

Industry	Constant	Coefficient	$\bar{\bar{R}}^2$
Public Utilities	267	-0.01 (0.004)	0.05
Distribution	2383	-0.1 (0.02)	0.33
Finance	1333	0.05 (0.02)	0.07
Public Administration	1074	0.07 (0.02)	0.10
TOTAL TERTIARY SECTOR	8342	0.44 (0.07)	0.32

To turn attention to the third difficulty posed by the original results, the above analysis shows that disaggregating the dependent variable does not improve upon the variance explained. In Table 7.4 the best results for each of the three independent variables in the final version of the model (equation (7.5)) are set out. When these are compared with the $\overline{R}^2$ obtained with the aggregated dependent variable (column 5) it is obvious that no independent variable explains more variance by itself for any industrial order than for the total.

Table 7.4

The most satisfactory results achieved with individual Industrial Orders as the Dependent Variable

Independent Variable	Highest results for individual industrial orders			$\overline{R}^2$ for total tertiary sector
	Constant	Coefficients	$\overline{R}^2$	
Large Town	1762	-0.004 (0.001)	0.13	0.20
Index of Earnings	2610	-3977 (754)	0.28	0.45
Export Sector	2383	-0.1 (0.02)	0.30	0.33

Furthermore, from Table 7.5, which gives the worst results achieved with the individual industrial orders, it follows that although all the independent variables performed adequately for some industrial orders none performed well for all, since each of the coefficients in column 3 is insignificant at the 95% confidence level. This may be taken as further evidence of the representative nature of the aggregated dependent variable since Table 7.6 shows that this form of the dependent variable does not underestimate the importance of any of the independent variables. Similarly, as shown in Appendix B Table B6, the distribution, finance and professional and scientific categories gave the best results, yet nevertheless the remaining industries were still important for certain independent variables. This latter point indicates that the overall result is not dependent simply upon one or two key industrial orders.

Table 7.5

The Least Satisfactory Results achieved with Individual Industrial Orders as the Dependent Variable

Independent Variable	Lowest Results for Individual Industrial Orders		
	Constant	Coefficient	$\bar{R}^2$
Large Town	211	-0.0002 (0.0002)	-0.01
Index of Earnings	1116	-2112 (1167)	0.05
Export Sector	267	-0.01 (0.006)	0.05

Finally, whilst still considering the individual industrial orders, the question to be asked is whether the full model [i.e. equation (7.5)], rather than the single independent variables, would yield a substantially different conclusion. The answer to this would appear to be in the negative, for when the individual industrial orders are substituted for the aggregate dependent variable in equation (7.5), the behaviour of the coefficients is not dissimilar i.e. they are either insignificant or are significant and carry the same sign as in equation (7.5). Moreover the best fit was achieved by the distribution order, yet again this explained a smaller percentage of the variance (50%) compared to the aggregate dependent variable (62%).

The second approach adopted in analysing the importance of various disaggreg ated versions of the dependent variable involved taking combinations of industrial orders. The method was to subtract successively from the total tertiary sector the individual industries in the reverse order of their importance as suggested in the first step above, and if the $\bar{R}^2$ declined with the omission of any one industry, then that industry was deemed to be an important component of the tertiary sector. The results obtained are set out in Appendix B Table B7 and can conveniently be summarised as follows:

(a) the industries that are not important appear to be public utilities, finance and the miscellaneous orders, though the results are not always consistent with step I above (see Professional and Scientific).

(b) the results can also be expressed in terms of which dependent variable the independent variable explains. From these tables it can be seen that the large town variable explains part of the variance in the transport, professional and scientific, public administration and distribution orders, the earnings variable the professional and scientific, the public administration and distribution orders and the export variable the transport, public administration and distribution orders.

In conclusion to the question of the aggregative nature of the dependent variable, although the results by themselves are interesting and throw considerable light upon the detailed behaviour of the individual industrial orders, they do not add significantly to the model as defined in terms of the total tertiary sector. Both the results for the industrial orders singly and in combination clearly show that no one industry behaves in a manner fundamentally different from the total tertiary sector, nor does any one independent variable appear to work consistently better in relation to other independent variables at the individual industrial order level. These conclusions follow from the facts that (a) no individual industrial order gives a better fit for the aggregative variable and (b) there was no disparity in the sign of the coefficients between the aggregative and individual industrial orders. However, it should be noted that these conclusions must be treated with caution, for even if the relationship was as strong at the individual order level, it would be expected that the resultant $\bar{R}^2$ would be lower due to the smaller variance to be explained at this level. In this sense the above negative conclusions, viz. that there is no evidence to suggest that the individual industrial orders behave in a fundamentally different manner from the aggregative variable, is preferable to the positive conclusion that the various hypotheses perform better for the aggregative variables.

Apart from investigating the possibilities of disaggregation, four other variations might be tried with the model in order to improve upon the overall variance explained. First, it has already been noted that the model ideally should be cast in terms of output rather than employment, but that the only output figures available were for the retail trade. However, when the output per employee is substituted as the dependent variable for distribution in employment terms, the $\bar{R}^2$ for the full model falls from 0.49 to 0.32. Although it is difficult to interpret a single result such as this, it would initially appear that the model is unlikely to be improved through a re-specification in output terms.

Secondly, there may be further scope for re-classifying the sample used. A classification using 38 and 55 towns has already been suggested upon the basis of an analysis of the behaviour of certain variables. One further division which might prove useful is that between the north and south of the country. Since the industries in the north are older and of a different character to those in the south this may lead to a systematic variation in the operation of the independent variables.

Details of all the regressions performed are set out in Appendix B Table B8 and the most important are reproduced below. From Table B8 it can be seen that there is a wide divergence in behaviour between the north and south samples, but that it is only those for the south which show any improvement over the full sample. Thus, for the latter, the large town variable gives the following equation:

$$S_t = 6533 - \underset{(0.007)}{0.03}\, L_t \qquad (7.13)$$

For this equation the $\bar{R}^2$ is 0.39 compared with an $\bar{R}^2$ of 0.20 for the full sample, thus illustrating that the large town variable has a greater explanatory power for the south sample as opposed to the full sample. The same conclusion is reached with the index of earnings variable in equation (7.14) below. This model has an $\bar{R}^2$ of 0.60 compared with 0.45 for the full sample.

$$S_t = 10289 - \underset{(3380)}{18675}\, T_t \qquad (7.14)$$

Finally, however, the fit of the export variable remains unchanged with an $\bar{R}^2$ of 0.32 i.e.

$$S_t = 9014 - \underset{(0.02)}{0.49}\, E_t \qquad (7.15)$$

When the large town and index of earnings variables are combined, equation (7.16) is formed:

$$S_t = 9647 - \underset{(0.006)}{0.02\,L_t} - \underset{(3224)}{14197\,Y_t} \qquad (7.16)$$

The coefficients both fall but still remain significant at the 95% confidence level. The most important point to note is that the two variables combine together to give a much higher $\overline{R}^2$ of 0.74 (compared to 0.51 for the full sample). However, the best specification was obtained, as with the full sample, using the large town, index of earnings and export sector variables, thus:

$$S_t = 10299 - \underset{(0.006)}{0.02\,L_t} - \underset{(3337)}{11875\,T_t} - \underset{(0.10)}{0.19\,E_t} \qquad (7.17)$$

For equation (7.17) an $\overline{R}^2$ of 0.74 was obtained.

To consider the above results, the large town variable explained 20% of the variance in the full sample, none in the north and 39% in the south. Several explanations may be put forward for this difference between the two samples. First in the south the towns developed at a later date and, since this coincided with a greater freedom of movement in the population, the towns were given ample opportunity to grow according to the respective influences of larger towns. In the north, on the other hand, the towns are older and when they were built population mobility was much lower so that they had to provide more of the services themselves, thus reducing considerably the possible influences of larger towns. The second explanation lies in the statistical difficulty of measuring the influence of larger towns since, in the north, the developments are so dense that the influence of one town upon another must inevitably become blurred. In the south (abstracting as the sample does from the influence of London) the population is spread more evenly, thereby facilitating the measurement of this variable.

Thus the difference in the behaviour of the large town variable is explicable in terms of the original hypothesis, but unfortunately this is not so with the index of earnings variable. From Table B8 this variable explains 60% of the variance in the south, but only 15% in the north (compared with 45% from the full sample). Yet, from an inspection of the residuals obtained with the full sample, there is no apparent significant difference between the samples. Unfortunately, therefore, any explanation of this difference must be delayed until the behaviour of this variable is further investigated in later chapters.

The third method by which the overall fit may be improved is to consider various transformations of the dependent and independent variable. Since the dependent variable is free to vary only between 0 and 100, but the independent variables are not so constrained, it might be expected that a combination of log transformations would be more satisfactory. However this proved not to be the case (Appendix B Table B9). Whether the independent variables in their normal or their log versions were regressed against the normal or log versions of the dependent variable, the results obtained were consistently lower and the coefficients less significant than for the normal versions of both the dependent and independent variables.

One final attempt to increase the overall variance explained lies in considering whether there are any towns which systematically exhibit extreme characteristics, for if this were so then the overall relationship could be weakened by the presence of towns which were influenced predominantly by local or other extraneous factors.

The residuals in the regressions of the five independent variables against the size of the tertiary sector 1961 were inspected and those which were significantly different at the 95% confidence level (i.e. which had values twice the standard error) were isolated. Table 7.6 below sets out those towns which were associated with such residuals for more than one of the five variables. However, considering the size of

Table 7.6

Towns with significantly different residuals

Towns	No. of variables for which residuals significantly different
Coventry	4
Scunthorpe	4
Chester	3
Sheffield	3
Stoke	3
Edinburgh	2
Kirkcaldy	2
Leicester	2
Southport	2

the full sample (69), the presence of such a few persistent extreme towns was not thought to be serious and, accordingly, it was not thought necessary to modify the sample to take account of these.

To summarise the findings on the attempts made in the above to improve upon the overall variance explained, it can be stated that, with the exception of the sample based upon the south of the country, they all gave less favourable performances than the original model. Thus re-specifying the dependent variable either in terms of output or in a disaggregated form does not increase the sophistication of the model. Neither does any combination of log transformations of the dependent and independent variables. Moreover, it would appear that there are relatively few persistent extreme towns so that these cannot be considered as a disruptive influence. As a result the only modification to be taken into the later chapters will be the division of the full sample upon a geographical basis. It would appear, therefore, that this last point apart, there is no reason to alter the model as summarised in equation (7.5).

The first version of the equilibrium model as given in equation (7.1) has now been studied in some detail and, with these results as a background, attention may now be diverted to the first differences version (i.e. equation (7.2)). The years upon which the relevant changes are based are 1951 and 1961, but unfortunately the use of the former is not without difficulty. As already noted in Chapter 6 the industrial structure was based upon the 1948 industrial classification, the main effect of which is to make the industries within the tertiary sector not strictly comparable with the 1958 classification system. For the purposes here, this has meant that the disaggregated versions of the dependent variable could not be investigated due to the presence of considerable intra-industry variation resulting simply from the difference in classification. Apart from this consideration, however, the procedure was the same as for the first version of the equilibrium model.

Unlike equation (7.1), the correlation coefficients proved unhelpful as a means of preliminary investigation since they were all too low, thus making it difficult to isolate the likely significant relationships. In order to discover more about these relationships, individual regressions were performed and all the resulting equations were found to be insignificant, with the exception of the index of earnings variable. Thus equations (7.18), (7.19), (7.20), (7.21) and (7.22) all have coefficients which are insignificant at the 95% confidence level. These equations relate to changes in the population, large town, small town, export sector and non-domestic values respectively.

$$\Delta S_t = 258 - \underset{(50.7)}{45.3}\,\Delta P_t \qquad \bar{R}^2 = -0.07 \qquad (7.18)$$

$$\Delta S_t = 221 + \underset{(0.01)}{0.007}\,\Delta L_t \qquad \bar{R}^2 = -0.01 \qquad (7.19)$$

$$\Delta S_t = 189 + \underset{(0.89)}{0.03}\,\Delta M_t \qquad \bar{R}^2 = -0.01 \qquad (7.20)$$

$$\Delta S_t = 212 - \underset{(0.03)}{0.005}\,\Delta E_t \qquad \bar{R}^2 = -0.02 \qquad (7.21)$$

$$\Delta S_t = 273 - \underset{(8.5)}{7.9}\,\Delta R_t \qquad \bar{R}^2 = -0.04 \qquad (7.22)$$

For the only significant variable, the $\bar{R}^2$ was 0.07 and the full specification was as follows:

$$\Delta S_t = \underset{(943)}{2123}\,\Delta T_t - 6.51 \qquad (7.23)$$

However, although equation (7.23) is significant, it is difficult to ascertain how much reliance should be placed upon the coefficient since, in this instance, it is positive. This might be thought to cast doubt upon the findings to equation (7.5), but it must be remembered that the index is based not only upon changes in actual earnings, but also upon changes due to the basis of classification. The presence of this bias in the results means that the coefficient must be treated with caution.

As with the first equilibrium version, the conclusions to the single regression equation remain unaltered when the variables are added in successively. Thus when the change in the index of earnings variable is combined with the change in the large town variable, the coefficient of the change in earnings variable is virtually unaltered, whilst the change in the large town variable is insignificant and the fit remains unaltered. Thus:

$$\Delta S_t = \underset{(948)}{2143}\Delta T_t + \underset{(0.01)}{0.008}\Delta L_t - 0.9 \tag{7.24}$$

Similarly, when the remaining significant variable in the first equilibrium model is added, the overall fit is unchanged and only the change in the index of earnings variable is significant at the 95% confidence level, i.e.

$$\Delta S_t = \underset{(956)}{2143}\Delta T_t + \underset{(0.01)}{0.009}\Delta L_t - \underset{(0.03)}{0.01}\Delta E_t - 4.2 \tag{7.25}$$

The pattern is repeated with the successive addition of the remaining variables, so that the final model has the form:

$$\Delta S_t = \underset{(896)}{2017}\Delta T_t + \underset{(0.01)}{0.0007}\Delta L_t - \underset{(0.03)}{0.009}\Delta E_t + \underset{(0.02)}{0.001}\Delta M_t - 4.9\Delta P_t - \underset{(8.5)}{6.9}\Delta R_t \tag{7.26}$$

Thus, although the coefficient upon the change in index earnings variable has fallen, it is still significant at the 95% confidence level, yet all the remaining variables fail to add any explanatory power. Moreover the sign of the index of earnings variable is still positive. It would appear that the change in population, large town, small town, export sector and non-domestic rateable values are all extraneous variables to this model and in no way augment the workings of the change in the index of earnings variable.

When the log transformations are tried (Appendix B Table B11) the sign is still positive, but the overall variance explained is reduced. A similar result is obtained for the 38 and 55 samples (Appendix B Table B12). With the north/south sample (Appendix B Table B13) that based upon the south is seen to perform worse than the full sample. For the north only the change in the large town variable gave a superior fit to the full sample i.e. for the north equation (7.19) was improved to:

$$\Delta S_t = 374 + \underset{(0.02)}{0.022}\,\Delta L_t \qquad (7.27)$$

For this equation the $\bar{R}^2$ was 0.05 but, despite the better fit, the coefficient was still insignificant at the 95% confidence level. When the variables are combined, the overall fit is improved with a combination of the change in the large town, index of earnings and export variables. Thus for equation (7.28) below an $\bar{R}^2$ of 0.12 was obtained:

$$\Delta S_t = 37.8 + \underset{(0.012)}{0.026}\,\Delta L_t + \underset{(1711)}{2714}\,\Delta T_t - \underset{(0.05)}{0.06}\,\Delta E_t \qquad (7.28)$$

Unfortunately, in this specification, the change in the export variable is insignificant. Yet when this variable is deleted, the overall fit (0.06) is worse than for the full sample, viz.

$$\Delta S_t = 40.6 + \underset{(0.02)}{0.024}\,\Delta L_t + \underset{(1560)}{2579}\,\Delta T_t \qquad (7.29)$$

It follows that the better fit for equation (7.28) is probably due to collinearity between ΔE_t and the other independent variables. Although in the final specification it is possible to have insignificant variables, this course should be avoided unless there are clear advantages to be gained through the inclusion of that variable. In this instance the improvement in the overall fit from 0.07 to 0.12 hardly makes this worth while, so the final specification for the first differences variable of the equilibrium model is given in terms of equation (7.23) i.e.

$$\Delta S_t = \underset{(943)}{2123}\,\Delta T_t - 6.51 \qquad (7.23)$$

DISEQUILIBRIUM MODEL

The above has been concerned solely with the equilibrium version. The disequilibrium model (equation (7.3)) takes as the dependent variable the difference in the percentages of employment in the tertiary sector in 1951 and 1961. From the a priori logic of Chapter 5 the independent variables could be based either on the 1951 or the 1961 data, and for this reason the regression analysis was conducted for both these years. The procedure in both cases was to repeat the regressions conducted for the first differences version of the equilibrium model, i.e. to take the model as it stands, to adopt the north, south, 38 and 55 samples, and to take different combinations of log transformations. However, since the dependent variable is the same as for equation (7.2), it was again not possible to conduct the analysis for the disaggregated versions.

1961 Data

The 1961 data is characterised by a complete lack of any relationship between the dependent and independent variables. Again in order to understand the form of the individual relationships, single regressions had to be performed and the following equations were obtained for the population, large town, small town, export, non-domestic rateable value and index of earnings variables respectively:

$$\Delta S_t = \underset{(74.9)}{61.5}\ P_t(\log) - 531 \qquad (7.30)$$

$$\Delta S_t = 145 + \underset{(0.001)}{0.001}\, L_t \tag{7.31}$$

$$\Delta S_t = 148 + \underset{(0.03)}{0.005}\, E_t \tag{7.32}$$

$$\Delta S_t = 48 + \underset{(0.04)}{0.02}\, M_t \tag{7.33}$$

$$\Delta S_t = 281 - \underset{(6.5)}{4.7}\, R_t \tag{7.34}$$

$$\Delta S_t = 1195 - \underset{(86.9)}{142}\, T_t \tag{7.35}$$

From an inspection of the coefficients and the standard errors, all the variables can be seen to be insignificant. In addition the $\bar{R}^2$ obtained for equations (7.30) to (7.35) were all negative (Appendix B Table B14).

No evidence to contradict the above conclusions is found when the variables are taken together. Thus if the large town and index of earnings variables are taken simultaneously equation (7.36) is obtained:

$$\Delta S_t = 197 - \underset{(98.5)}{131}\, T_t + \underset{(0.01)}{0.001}\, L_t \tag{7.36}$$

In this equation both the coefficients are insignificant and the $\bar{R}^2$ is negative. As equations (7.37) to (7.40) below indicate, the same conclusions can be drawn from the successive addition of the remaining variables:

$$\Delta S_t = 216 - \underset{(970)}{127}\, T_t - \underset{(0.01)}{0.001}\, L_t - \underset{(0.02)}{0.003}\, E_t \tag{7.37}$$

and, with the addition of the population variable,

$$\Delta S_t = \underset{(85.6)}{50.9}\, P_t(\log) - \underset{(92.7)}{109}\, T_t - \underset{(0.005)}{0.0007}\, L_t - \underset{(0.008)}{0.009}\, E_t - 114 \tag{7.38}$$

Similarly the inclusion of the small town variable adds nothing to the model, viz.

$$\Delta S_t = \underset{(80.9)}{55.6P_t}(\log) - \underset{(108)}{109T_t} - \underset{(0.003)}{0.007L_t} - \underset{(0.008)}{0.009E_t} + \underset{(0.03)}{0.02M_t} - 94 \quad (7.39)$$

Finally, a similar result is obtained with the non-domestic rateable value variable:

$$\Delta S_t = \underset{(78.5)}{49.5P_t}(\log) - \underset{(96.3)}{123T_t} - \underset{(0.005)}{0.007L_t} - \underset{(0.007)}{0.006E_t} + \underset{(0.03)}{0.02M_t} + \underset{(7.6)}{6.8R_t} - 208 \quad (7.40)$$

<u>1951 Data</u>

The main difference between the 1951 and 1961 data is that the index of earnings variable explains 24% of the variance. Thus, for 1951, equation (7.35) is altered to:

$$\nabla S_t^{1} = 2938 - \underset{(4482)}{18829\ T_t} \quad (7.41)$$

However, with the addition of the remaining variables the coefficient upon the index of earnings variable remains virtually unchanged. Thus if the large town variable is considered alongside this variable the model is specified thus:

$$\nabla S_t = 2918 - \underset{(4653)}{18717\ T_t} + \underset{(0.001)}{0.0001\ L_t} \quad (7.42)$$

Also, with the addition of the last remaining significant variable in the equilibrium model, the export sector variable, the coefficient on the index of earnings variable is still significant and the others insignificant at the 95% confidence level:

$$\nabla S_t = 3017 - \underset{(4734)}{18930\ T_t} + \underset{(0.0014)}{0.0001\ L_t} - \underset{(0.03)}{0.01\ E_t} \quad (7.43)$$

Finally, to consider different samples (Appendix B Table B15 and B16) or

---

1. The notation ∇S_t refers to the change $\left[S_{(t+1)} - S_t\right]$ i.e. the change using 1951 data.

to take log transformations of either the dependent or independent variables (Table B17), fails to improve upon the above results.

Thus, in conclusion to the disequilibrium model, for 1961 there is a complete absence of any relationship, whilst for 1951 the model would be specified as in equation (7.41). In this equation the index of earnings variable carries a negative sign as in the first version of the equilibrium model (but in contrast to the first difference model). Finally, from the behaviour of the coefficients and standard errors of the remaining variables, it is fair to conclude that they are extraneous information variables.

CONCLUSIONS

The two versions of the equilibrium model may be specified as follows:

$$S_t = 10529 - \underset{(0.003)}{0.013}\,L_t - \underset{(2228)}{11581}\,T_t - \underset{(0.06)}{0.27}\,E_t \qquad (7.5)$$

and $$\Delta S_t = \underset{(943)}{2123}\,\Delta T_t - 6.51 \qquad (7.23)$$

Equation (7.5) differs from the original hypothesis as specified in equation (7.1) in two ways. First the hypotheses as represented by the population, small town and non-domestic rateable value variables have been found to be insignificant and the variables accordingly deleted. Secondly, the index of earnings and export sector variables carry a negative coefficient thereby contradicting traditional economic theory. The first differences model (equation (7.23)) has been found to be a far less satisfactory version of the equilibrium hypothesis in the sense that it contains only one variable and explains only 7% of the variance.

However it should be noted that for this version the income variable carries a positive coefficient.

The disequilibrium model is specified in equation (7.41) above. This model can only show a relationship for the 1951 data, but for that year it does give the better fit of the two versions of the model which take as the dependent variable the change in the tertiary sector. Moreover, this model would appear to reinforce the finding of equation (7.5) with respect to the negative coefficient of the index of earnings variable.

Finally, various more sophisticated forms of the basic models have been tried, but with the exception of the north/south version of equation (7.23) they fail to bring about any increase in the variance explained.

Chapter Eight

THE MODEL: SOME EXTENSIONS

In Chapter 5 a model, as deduced from a survey of all the literature in the field, was postulated which took as its basic variables population, geographical location or function, the size of the exporting sector and the level of per capita income. In the last chapter it was seen that, with reservations, population failed as an explanatory variable and that, of the two function variables, the "small town" variable was also ineffective. Furthermore, the last chapter explored the possibility that the aggregative nature of the dependent variable masked the true behaviour of the independent variables but, although the results of the various disaggregated versions tried were in themselves interesting, they collectively failed to bear out this proposition. Finally, in the light of the surprising results obtained for the export and the index of earnings variable their behaviour and similarities were further investigated. From this the conclusion was drawn (albeit tentatively) that both variables were measuring the same influence, namely some aspect of a town's per capita income (the possibility that the specification of the variables was the main cause of the negative coefficient having been largely discounted).

It would appear from the results of the last chapter that, apart from the geographical function variables, either traditional theory is irrelevant or it is simply incorrect when applied at the town level. If this is so then the next stage in this analysis must be to discover what other factors are likely to influence the size of the tertiary sector in towns.

It should be noted at this stage that although the results of the last chapter cast doubt upon traditional theory, the results themselves may still be used as a base for further inquiry. In essence the last chapter, using 'traditional' variables, managed to explain 62% of the variance for the equilibrium model, but only through the use of two variables with coefficients carrying the 'wrong' sign. Clearly one aim of any further analysis must be to explain this 'wrong' sign. Nevertheless a further aim should also be to look beyond these variables and to discover new factors which will increase the total variance explained. It must be realised that the model in Chapter 5 could only be deduced rigorously from previous theoretical writings on the subject, but that it is now possible to put forward theoretical formulations, based on the experience of the last chapter, which may be more directly relevant to the precise problem at hand.

Finally, a third way in which the results already obtained may be improved is to consider those variables which have already been found to be important and to consider in what ways they might be refined.

Consequently this chapter will be concerned with extending the model deduced in Chapter 5 in an attempt to make it more relevant to a study of inter-town differences in industrial structure. This aim will be approached through a) an investigation of the cause of the negative coefficient on the earning and export variables, b) a formulation of new hypotheses in the light of the results of the last chapter and c) an investigation of possible sources of improvement in those variables already shown to be important and significant.

NEGATIVE COEFFICIENT ON EXPORT SECTOR AND INCOME VARIABLES

To consider the export variable first, one explanation may be due to the fact that towns which export a large proportion of their output are also, typically, towns with a large secondary sector. Thus although the increased flow of income resulting from a higher export activity would tend to enhance the size of the tertiary sector, these towns would still appear to have below average tertiary sectors in a cross-section study. Such an association may be explicable in historical terms. Often towns in the early part of the century expanded predominantly under the influence of a single industry. This lent to the town an unbalanced industrial structure the chief aspect of which was a large secondary sector. Examples of such towns would be the textile industry towns in Lancashire, the woollen towns in Yorkshire, the ship-building towns in the North-east, etc.

If this hypothesis is correct then what variables may be used to measure this phenomenon? Obviously one direct measure would be provided by an index of the extent to which a town's secondary sector in 1961 was concentrated upon those industries important at the turn of the century. If this index is represented by (C) then the hypothesis becomes:

$$S_t = F(C_t) \qquad (8.1)$$

where S_t = the size of the tertiary sector 1961.

This hypothesis could be checked by seeing whether the export variable (E) was itself a function of this 'index of concentration' i.e. whether

$$E_t = F(C_t) \qquad (8.2)$$

An alternative method would be to consider the consequences of a town with the above described industrial structure. It may be postulated that those towns which relied heavily upon a single industry would be unlikely to exhibit a fast rate of population growth for several reasons. First, most of these industries have been declining. Secondly, since these towns possessed only one industry and, therefore, a specialised labour force, the new, expanding and typically foot-loose industries would not be attracted to these towns, thus reducing further the possibility of a population expansion. It would therefore not be unreasonable to expect a low population growth to be associated with the extent to which a town's labour force was concentrated in either one or more industries which were important at the turn of the century. If the original hypothesis is correct, namely that these towns are associated with large secondary sectors, then <u>ceteris paribus</u> towns with a low population growth will be those with small tertiary sectors. Moreover the corollary to a low population growth, connected as it is with declining industries and limited employment opportunities, would be a low income, so that this relationship would be further reinforced.

Thus the alternative hypothesis may be expressed as:

$$S_t = F(P_T - P_{(T-1)}) \qquad (8.3)$$

where P_T = population in 1961

$P_{(T-1)}$ = population in a selected year.

As already noted elsewhere, the export variable only measures in a rudimentary fashion the extent to which a town exports and as a result of the aggregativeness of its construction the possibility that the variable was measuring something quite different was explored in the above two hypotheses. These hypotheses may be considered as arising solely from the statistical difficulties of the export variable. This point apart, it has already been noted that a tentative conclusion to the last chapter was that both the export and the index of earnings variables in fact reflected per capita income. Consequently further explanations of the negative coefficient may be sought through considering the relationship between per capita income and the tertiary sector and the following remarks may therefore be taken as applying to the index of earnings variable as well as the export variable.

If the income of a town increases then the effect upon the tertiary sector will be transmitted via the pattern and/or volume of demand, and it is therefore reasonable to look for an explanation of the negative coefficients in the behaviour of demand. It is suggested here that this will lead to three separate hypotheses, not all of which are capable of providing a complete explanation by themselves.

The size of the tertiary sector has been calculated in terms of employment, yet the effect upon demand is in terms of output. Clearly, therefore, labour productivity must be important in determining the nature of the relationship between income and the size of the tertiary sector. It could be legitimately presumed that an increase in demand
required will bring forth a rise in productivity and, although the example of the super-market is obvious, there is no reason why this process should be confined to distribution.

Although changes in productivity could never explain an actual decline in the size of the tertiary sector as a result of the rise in per capita income, it must clearly cause a considerable decrease in the rate of expansion of the tertiary sector, whilst the undoubted indivisibilities in the production function may well lead to virtually no increase in employment. For instance an initial rise in demand may lead to the provision of a new super-market and, thereafter, the increase in employment would be small over a considerable range of further demand increases.

The above argument may be expressed in the hypothesis that:

$$S_t = F(X_t) \tag{8.4}$$

where X_t = output per employee.

As already stated this hypothesis cannot explain any decline in the tertiary sector (i.e. the negative coefficient itself), but it may neutralise the effect of an increase in demand, in that further factors would require a smaller impact in order to produce a downward sloping relationship.

The above approach concentrates upon aggregate demand, but a further insight may be gained through a consideration of the actual pattern of expenditure in towns. In this context it should be noted that the services may either be supplied by the towns themselves or imported from elsewhere. Consequently, upon a rise in per capita income the resultant increase in income could be met through either an expansion of 'home-produced' services or imports. Thus, for instance, a rise in income may lead to more being spent upon the local cinema or people may travel to a theatre in another town. An effect similar to the latter would result where, although the actual purchase was made in the sample town, the provision of the service was basically left to that of another e.g. most of the trade transacted through local agents and branches. It follows, therefore, that if the pattern of demand should shift towards imported services upon a rise in income, the expected increase in the size of the tertiary sector would be curtailed. Moreover such a shift would be intelligible since sophisticated service consumption would rise faster than that of the essential services. Nevertheless, as with productivity increases, this effect by itself could not lead to an explanation of the actual decline in the tertiary sector.

This hypothesis is similar to the 'large town' concept, namely that expenditure may be directed from one town to another. But here it is the income of the sample town, irrespective of the presence of others, which causes the change in demand, whilst in the geographical variable it is the presence of the external influence of other towns that is important.

Thus the general hypothesis may be expressed as:

$$S_t = F\left[K_n - K_I\right] \tag{8.5}$$

where K_n = the demand for services supplied within the town

K_I = the demand for services imported into the town.

Unfortunately there is no direct way of measuring the imported content of a town's services consumption and it is necessary to take the argument further in order to arrive at a testable hypothesis. If the services are imported into the town then they must be supplied from somewhere else, and the natural presumption is that it would be the large recognised service centres offering a wide range of sophisticated services which would be the main providers. In addition it could be reasonably presumed that the services of these regional centres would be supplied over the immediate area around them rather than on a national scale. As a result the demand facing the larger service centres themselves would depend partly upon their own per capita income, but also upon the per capita income of the service hinterland. In other words, if the income of the sample towns in the hinterland is above average, then a greater proportion of the demand will be for imported services and hence the size of the tertiary sector of the associated service centre will be larger than that indicated by its own per capita income, i.e. that:

$$S_c = F\left[Y_c, Y_H\right] \tag{8.6}$$

where S_c = the size of the tertiary sector of the service centre

Y_c = the per capita income of the service centre

Y_H = the per capita income of the hinterland.

So far the pattern of demand has been analysed only across the broad division between imports and home-produced services, the actual services purchased having been ignored. Yet an analysis of actually what type of services are purchased must clearly play an important part in an understanding of the income variable. From the Family Expenditure Survey (H.M.S.O.) it is possible to discover the percentage changes in expenditure over a wide range of consumption and these are reproduced below in Table 8.1 for the period 1962-66. From this it is clear that, apart from transport, it is only expenditure on housing that rises with per capita income. Moreover it would appear from the behaviour of both the mortgage and the own dwelling classifications that it is private

Table 8.1

Changes in Consumers' Expenditure Pattern

Period 1962-1966

Variable	Percentage Change
TOTAL INCOME	+ 30.9
Housing	+ 7.7
Own dwelling	+ 29.7
Fuel	- 8.4
Food	- 11.4
Drinks and Tobacco	- 2.4
Clothing	- 8.0
Household durables	- 5.2
Other goods	- 8.4
Transport and vehicles	+ 4.6
Services	- 6.4
Mortgages	+ 13.0
Other savings	- 2.9

housing which shows the greatest rise. Upon the basis of this information it might be expected that ceteris paribus towns with a higher per capita income would spend a larger percentage of their income upon private housing. Some of this expenditure may be merely reflected in higher prices for existing houses, but part of it must result in a greater demand facing the construction industry, either in the form of new houses or in improvements made to the existing capital stock. Such a shift in the pattern of demand would be significant because the definition of the tertiary sector adopted in this study has excluded the construction industry, so that this may contribute further to the undermining of the traditonal relationship.

However to follow the procedure above and to take that to which the demand is diverted as the independent variable would be inaccurate in this case (since the construction industry cannot be disaggregated), and the hypothesis might therefore be masked by the inclusion of activities other than house-building. For this reason the more direct formulation that the size of the construction industry might be explicable in terms of the per capita income has been adopted here, viz:

$$C_t \; = \; F\left[Y_t\right] \qquad (8.7)$$

where C_t = the size of the construction industry

Y_t = the per capita income of a town.

To summarise this section, in the absence of any real suggestion that the variables have been mis-specified [Appendix E], an explanation for the negative coefficient in the export and earning variables has been sought in a) the statistical weaknesses of the export variable and b) the relationship between per capita income and the pattern of demand within towns. The former suggested that either population change or the extent to which a town's labour force was concentrated in certain industries might themselves be independent variables in the model as deduced in Chapter 5. The second approach has yielded certain influences, namely productivity changes, the consumption of imported services and the expenditure on housing which, although by themselves ~~they~~ could not lead to an explanation of the negative coefficient, would nevertheless modify or reduce the traditional relationship.

FORMULATION OF NEW HYPOTHESES

The model deduced in Chapter 5 followed logically from a review of the literature and all the main suggestions put forward in these writings have been included in the model. As a result purely theoretical a priori reasoning can be ruled out as a possible source of further formulations. On the other hand, the results of the last chapter have thrown some light upon how towns actually behave which suggests that this knowledge might now be used to specify new hypotheses.

One finding of the last chapter which suggests itself as the starting-point for further inquiry is the wide divergence in the size of the tertiary sector in towns of a similar nature. Take for instance Chester (pop. 59,268) and Swindon (pop. 91,159). From the location of these towns it would be expected that both would perform similar functions i.e. both represent the main towns in the immediate area yet both are within easy access of very large towns. Moreover the index of earnings are similar being 0.271 for Swindon and 0.232 for Chester. From the large town variable it would be expected that Swindon with a value of 17623 would have a larger tertiary sector than Chester with a value of 74441, yet the size of the tertiary sector for Chester is 74% whilst for Swindon it is 51%, i.e. 32% lower. It may therefore be legitimately asked why towns so similar in respect to the variables tested in the last chapter exhibit such a wide divergence in industrial structure?

The last chapter showed that theoretical considerations can explain a considerable amount of the variance in towns, yet there still appear to be further elements unexplained. One aspect of the problem which has so far been ignored is the question of a town's historical development. The size of the tertiary sector in towns undoubtedly is influenced strongly by its geographical location, income etc., but part of its development must at least be due to historical circumstances peculiar to the towns themselves. Thus, although the size of the tertiary sector would be influenced by the variables so far specified in the model, the model could not have been expected to reach at the full explanation due to the presence of this historical element.

An example of this hypothesis might be a town in which its industries required a large amount of services or which had become used to a wide choice of services. More likely, perhaps, the town acquired the reputation of a service centre so that new service industries starting in the area, or existing services expanding, would naturally prefer to locate in that town. In addition consumers would automatically tend to go to that town for their services irrespective of whether they could be provided elsewhere. Thus, for instance, Chester may have a reputation as a service centre, whilst Swindon has not, so that the provision of services would then be a characteristic of Chester, which is a point which so far could not have been picked up by the model.

This, in effect, implies that the previously tested variables are important, but that superimposed upon these should be some consideration of the town's reputation or characteristic as a service centre. Thus, if the town had an above average sized tertiary sector some years ago as a result of this historical factor, it will still be likely to have one in 1961. In other words the size of the tertiary sector in 1961 is partially dictated by what it was some time ago i.e:

$$S_t = F\left[S_{(t-1)}\right] \qquad (8.8)$$

where S_t = the size of the tertiary sector 1961

$S_{(t-1)}$ = the size of the tertiary sector in a selected year.

The hypothesis in equation (8.8) measures indirectly the extent to which the historical development of a town has influenced its industrial structure. However, if it is the traditional attractiveness of a town as a tertiary sector that is the main influence behind this equation, the hypothesis may be measured more directly through considering the likely manifestations of that attractiveness. If a town has for some time traditionally been a centre for the provision of services this would imply that it was also the focal point of the region. Furthermore one obvious indicator of the extent to which a town acts as such a centre is provided by the number of hotels it has to offer, for clearly if it is the focal point of the area it will be the point to which people will naturally travel. Moreover, apart from this convenience point, such towns will also be more attractive to the traveller either physically due to the attendant lack of heavy industry or through the range of ancillary services offered. Thus to take the example of Chester again, travellers in the area are likely to be attracted to that town because it is the focal point of the area, it has the tradition of being a service centre and it is likely to be able to offer a greater range of services. It would therefore be not unreasonable to place such a town in a category different from that of a town without such a character say, for instance, Barnsley.

Hence the above suggestions lead to the hypothesis that:

$$S_t = F[H_t] \qquad (8.9)$$

where H_t = the number of hotels in the town.

To summarise this section, the model in Chapter 5 took into consideration all the theoretical formulations available so that any new hypothesis would have to stem from the practical experience gained in the last chapter. One of the findings of that chapter was that towns, even although similar with respect to the independent variable, differed considerably in the size of their tertiary sector, thus suggesting that some influence had as yet been unidentified. It is suggested here that this may be represented by the historical development of the town and this may either be measured by the extent to which the tertiary sector now is explicable in terms of what it was a number of years ago, or, more directly, by the extent to which hotel accommodation is provided within the town.

REFINEMENTS OF THE EXISTING VARIABLES

The variables which were found to be important in the last chapter were the large town, export and income variables and the question now raised is whether, by considering them further, they may be improved upon. In the first section of this chapter the export and income variables were investigated and in trying to explain the negative coefficients any available refinements have necessarily been introduced. Thus the only variable remaining to be considered here is the geographical location variable.

It will be recalled that this variable attempted to define a town's relationship with the area around it. Although the data on the various relationships which a town has with its neighbours were not available, nevertheless the net effect of the influence of one town upon another was measured through the use of the gravity formula. Two modifications might suggest themselves. First, either the gravity model could be improved as a measure of the interaction between towns or additional considerations may be introduced to supplement the workings of this model. In Chapter 5 it was indicated that the main source of error in the gravity formula lay in the values to be assigned to the 'weights', yet this was a subject explored in some depth in the last chapter. It would appear that the possibilities of further improvements in the gravity formula are accordingly limited. However, in defining the relationships between the sample town and its hinterland, the gravity formula omits one important aspect of a town's geographical location, namely the extent to which people travel to the town to work. If, as a result of the location of the town, an above average percentage of the labour force commutes from outside the town's boundaries, then this is likely to have important consequences for the size of the tertiary sector. However, on a priori grounds alone, this would either result in an expansion or a contraction of that sector. Thus if people commute from outside, then part of their income will be spent in that town which otherwise would have been spent elsewhere, and this would lead to an expansion of the tertiary sector. Alternatively, it could be argued that if some people travel to work from outside instead of living within the town's boundaries, then part of their expenditure will be lost to the sample town. In either case the geographical location variable ought to be modified or supplemented by the inclusion of some 'index of commuting' i.e. that the following hypothesis is likely to be relevant:

$$S_t = F\left[A_t\right] \qquad (8.10)$$

where A_t = index of commuting.

So far the model has been specified with the dependent variable expressed as a percentage of the total labour force. One further suggestion might be to consider the dependent variable in absolute terms. Although a large part of the employment in the tertiary sector must of necessity be due to the town's population, there may nevertheless be sufficient variance remaining to make this avenue of enquiry worth while. That being the case, the attention may be diverted to either the model in both its equilibrium and disequilibrium forms.

SUMMARY

In the last chapter much of traditional theory was shown to be either irrelevant or incorrect and it has been the task of this section to extend the model of Chapter 5 in order to make it more realistic. Three separate starting-points of trying to explain the negative coefficient on the export and income variables, of formulating new hypotheses in the light of the knowledge gained in the last chapter and of considering further the variables found to be relevant were taken and these have resulted in the hypotheses spelt out in equations (8.1) to (8.10). In the following chapter these will be adjusted so as to be consistent with the available data, tested and those found to be worthwhile will be integrated with the results of Chapter 5.

Chapter Nine

THE RESULTS II

The task of this chapter is two-fold. First, the individual hypotheses of the last chapter will be tested in accordance with the available data and, secondly, those hypotheses found to be significant will be integrated with the main conclusions to the original model.

INDIVIDUAL HYPOTHESES

Six of the hypotheses put forward in the last chapter took the percentage size of the tertiary sector as the dependent variable and these may be presented thus:

$$S_t = F\left[C_t\right] \qquad (9.1)$$

where S_t = the size of the tertiary sector and C_t = an index of concentration of employment in the older industries.

$$S_t = F\left[P_t - P_{(t-1)}\right] \qquad (9.2)$$

where $\left[P_t - P_{(t-1)}\right]$ = the change in population from a selected year.

$$S_t = F\left[X_t\right] \qquad (9.3)$$

where X_t = output per employee.

$$S_t = F\left[S_{(t-1)}\right] \qquad (9.4)$$

where $S_{(t-1)}$ = the size of the tertiary sector in a selected year.

$$S_t = F\left[H_t\right] \qquad (9.5)$$

where H_t = index of hotel accommodation.

$$S_t = F\left[A_t\right] \qquad (9.6)$$

where A_t = index of commuting.

In addition equation (9.1) above may be alternatively tested by considering the related hypothesis that -

$$E_t = F\left[C_t\right] \quad (9.7)$$

where E_t = the size of the exporting sector.

Further, the two remaining relationships put forward in the last chapter were -

$$S_c = F\left[Y_c, Y_H\right] \quad (9.8)$$

which related the size of the tertiary sector of service centres (S_c) to their per capita income together with an index based upon the per capita income of their hinterland, and

$$Z_t = F\left[Y_t\right] \quad (9.9)$$

where Z_t = the size of the construction industry and Y_t = the per capita income of the town.

Finally (9.9) may be specified in terms of domestic rateable value (D_t), i.e.

$$D_t = F\left[Y_t\right] \quad (9.10)$$

Before these hypotheses can be used as the basis for a regression analysis they must be both defined more precisely and made more consistent with the available data. In specifying the 'index of concentration in older industries' (hereafter called simply the 'index of concentration'), those industries which were important at the turn of the century were taken as metal manufacture, metal goods not elsewhere specified, shipbuilding and marine engineering and textile industrial orders (the engineering and electrical goods category was not considered since it included too many modern industries). The index was then compiled by dividing the aggregate percentage employment in these industrial orders by the percentage size of the secondary sector (all percentages relating to the total labour force).

In equation (9.2) the percentage change in population was calculated between 1961 and 1901, 1911, 1921, 1931 and 1951 (Appendix A Table A5), since from the logic of this hypothesis the relationship could either be a short-term or a long-term one. For equation (9.3) it has already been noted in the last chapter that the effect of productivity changes would not be confined simply to distribution, but unfortunately it is only for this industrial order that output data is available. In the census of distribution both total and full-time employment figures are given, which allowed the opportunity to construct two indices for productivity, one including and one excluding part-time employment.

The rationale to equation (9.4) rests upon towns having a certain characteristic or reputation for being a service centre. This of necessity is an influence that has to be measured over a considerable period of time and for this reason the year 1931 was adopted.[1] Equation (9.5) measures the same influence more directly through the provision of hotel accommodation. This information was taken from the A.A.Handbook and the index was computed by first determining the average number of hotels for selected population intervals, and, secondly, taking the difference between this and the actual numbers provided by each town.

1. Unfortunately the census of population presents data on industrial structure only for those towns with populations in excess of 50,000. Since some of the sample towns in 1931 had populations less than this figure, data on this variable is not obtainable and the samples based upon this variable are accordingly restricted. The towns for which the data was not available were Bedford, Cambridge, Colchester, Crewe, Harrogate, Mansfield, Nuneaton, Peterborough, Rugby, Scunthorpe, Torquay, Widnes, and Worthing, giving a sample of 55.

In the last chapter the index of commuting was intended to reflect the percentage that the town's labour force bore to its population and, since both figures are given in the Census of Population, no difficulty was experienced in specifying this variable.

Of the four relationships (9.7) to (9.10) which did not take the size of the tertiary sector as the dependent variable, only equation (9.8) presented any specification problems. The main service centres in the sample of 69 towns were taken as the chief regional centres in the sample viz. Bristol, Cardiff, Edinburgh, Hull, Plymouth, Sheffield, Stoke and Nottingham and their hinterlands were defined as those sample towns within a radius of 50 miles. The index which reflected the income of their hinterland (Y_H) was calculated in three ways. First, a simple average of the index of earnings of the secondary sectors of the hinterland towns was adopted. Secondly, recognition was given to the actual amount of purchasing power generated by these towns through calculating the actual wage bill of each town (based upon the employment in each industry and the national wage earnings for that industry). In addition it might be thought that the influence of distance would be important, so the purchasing power thus derived for each town was divided by its distance from the service centre, thereby generating a third index.

The results obtained when these hypotheses are evaluated using 1961 data are set out in Appendix B Tables B18 and B19, and the full equations are reproduced below. For equation (9.1) the full specification was

$$S_t = 6365 - \underset{(758)}{3466}\, C_t \qquad (9.11)$$

The coefficient is significant at the 95% confidence level and with an $\overline{R}^2$ of 0.20 the result is not inconsistent with the hypothesis put forward in equation (9.1). However, with the alternative version of this model, equation (9.7), the relationship was found to be insignificant at the 95% confidence level and the $\overline{R}^2$ obtained was only 0.09.

From Appendix B Table B18 it can be concluded that the various population changes in equation (9.2) all fail to provide a satisfactory relationship. It would appear that the change 1921 - 1961 offered the best possibility, yet as can be seen from equation (9.12) the relationship is still insignificant at the 95% confidence level.

$$S_t = 5385 + \underset{(8.7)}{14.0} \Delta P_{1921} \qquad (9.12)$$

Similarly the hypothesis as embodied in equation (9.3) was found to be not substantiated. The version based upon the output per employee gave the highest $\overline{R}^2$ (0.05), but the coefficient is still negative, e.g.

$$S_t = 6706 - \underset{(30.8)}{40.9} X_t \qquad (9.13)$$

On the other hand, the hypothesis embodied in the percentage size of the tertiary sector 1931 (equation (9.4)) gave a most pronounced fit with an $\overline{R}^2$ of 0.74. The model was specified as -

$$S_t = 1432 + \underset{(0.07)}{0.84} S_{1931} \qquad (9.14)$$

It follows from equation (9.14) that the size of the tertiary sector is markedly influenced by the 'historical' element as represented by the past size of the tertiary sector. Moreover, the alternative method of measuring this influence, using an index of hotel accommodation (equation (9.5)), also gave a significant relationship, though this time the $\overline{R}^2$ fell to 0.25.

$$S_t = 5261 + \underset{(25.2)}{112}\, H_t \qquad (9.15)$$

Thus equations (9.14) and (9.15) clearly point to the 'historical element' as having a powerful influence upon the dependent variable.

The final hypothesis taking the dependent variable as the size of the tertiary sector also gave a significant relationship. Thus the index of commuting gave rise to the equation -

$$S_t = 12344 - \underset{(41.1)}{146.7}\, A_t \qquad (9.16)$$

The $\bar{R}^2$ is rather low (0.17), but the negative coefficient does indicate that of the two possible effects of this variable (Chapter 8) it is the argument based upon a net loss of income to the town that is preponderant.

When the remaining hypotheses were tested, it was found that the inclusion of the hinterland (equation (9.8)) failed to explain any more variation in the size of the tertiary sector over and above that already explained by the index of earnings alone. Thus, for these 'centres', the income variable gave rise to the equation -

$$S_c = 16375 - \underset{(15826)}{44316}\, T_t \qquad (9.17)$$

for which the $\bar{R}^2$ was 0.49 and the coefficient significant at the 95% confidence level. However, the inclusion of the indices of earnings in the hinterland (Y_1 to Y_3) failed to improve upon the model as given in equation (9.17). Thus the inclusion of the index representing the average of the earnings in the hinterland caused the $\bar{R}^2$ to fall to 0.45 and its coefficient was insignificant, i.e.

$$S_c = 15154 - \underset{(19750)}{36611}\, T_t - \underset{(4478)}{3188}\, Y_I \qquad (9.18)$$

The same result was achieved with the index reflecting the purchasing power of the hinterland (Y_2) and the index weighted for distance (Y_3)

$$S_c = 23555 - \underset{(40142)}{76172}\, T_t + \underset{(0.04)}{0.03}\, Y_2 \qquad (9.19)$$

and

$$S_c = 18152 - \underset{(31759)}{52278}\, T_t - \underset{(0.77)}{0.22}\, Y_3 \qquad (9.20)$$

The $\overline{R}^2$'s for equations (9.19) and (9.20) were 0.47 and 0.40 respectively. It follows from equations (9.17) to (9.20) that the inclusion of the hinterland fails to improve upon the model and the continuing significance of the index of earnings variable illustrates that it provides in this instance the main explanatory power. Hence the hypothesis that an increase in income leads to a shift in the demand towards imported services is not substantiated, since there does not appear to be any increase in the size of the tertiary sectors towards which this demand might be expected to be directed.

The final hypothesis to be tested was that the rise in income might lead to a rise in the demand for housing as reflected either in the demand for construction (equation (9.9)) or domestic rateable value (equation (9.10)). Unfortunately both these hypotheses were found to be incorrect. For the construction industry (Z), the equation was -

$$Z_t = 1189 - \underset{(62.5)}{2032}\, T_t \qquad (9.21)$$

The fit was very pronounced with an $\overline{R}^2$ of 0.41, but with a negative coefficient this illustrates that, in this respect, the construction industry is simply an extension of the tertiary sector. Had there been a switch of demand into the construction industry, as hypothesis (9.9) would suggest, the coefficient would have been positive. This point is finally confirmed by the behaviour of equation (9.10), for when it is estimated the sign is negative, the coefficient significant at the 95% confidence level and the $\bar{R}^2$ is 0.25 -

$$D_t = 13.7 - \underset{(7.2)}{15.3}\, T_t \qquad (9.22)$$

Conclusions

The conclusion to be taken from the individual hypotheses is that, on the above evidence, the indices for commuting, concentration and hotel accommodation, together with the percentage size of the tertiary sector in 1931, should be incorporated into the main findings of the original model.

However, the conclusions go further than this for it will be recalled from the last chapter that the hypotheses themselves were formulated specifically to -

(1) explain the negative coefficient on the 'income' variables,

(2) present further hypotheses to be tested, and,

(3) to provide more sophisticated versions of the original variables in the theoretical model.

Since there is no direct measure of income at the county level, the theoretical formulations of Chapter 5 have to be expressed indirectly, viz. the theoretical formulations are supposed to be:

Index of earnings ⟶ level of income ⟶ size of tertiary sector
Export sector ⟶ level of income

Apart from the possibility of a mis-specification of these variables (which has been largely discounted), the explanation for the negative coefficient has been sought in the connection between either the proxy variables and income or income and the size of the tertiary sector.

To consider the first relationship, it was one of the findings of Chapter 7 that the index of earnings was an acceptable proxy variable for income, but in the last chapter the statistical difficulties associated with the export sector variable cast doubt upon the validity of its connection with the level of income. This was postulated as being the result of the tendency for towns with large export sectors to be those which developed in the early part of this century. This meant that the above relationships could be destroyed either directly due to these towns being associated with large secondary sectors (i.e. by-passing the link through the level of income) or indirectly due to these towns having a lower per capita income. The latter point was held to be likely in view of the declining nature of these industries together with the relative unattractiveness of the towns to the footloose and high income earning industries. In the above empirical work, the direct relationship was reflected by the index of concentration and the indirect by the change of population.

The results as a whole dispute the above arguments. Certainly the failure of the population change variable suggests that the indirect relationship is incorrect, whilst an $\bar{R}^2$ of 0.20 obtained with the index of concentration is not itself conclusive. The result is significant, indicating that the more the towns specialise in those industries which were important at the turn of the century the smaller is their tertiary sector, but if the real cause of the negative coefficient on the export sector variable is to be due to this explanation then that variable must be measuring this index of concentration. Yet when equation (9.7) is evaluated, this is found not to be the case.

Thus with regards to the first connection between the proxy variables and the level of income, the preliminary investigations in Chapter 7 suggested that the index of earnings variable is related to the level of income. The main difficulty in this section is the connection between the export sector variable and the level of income, for at the least this must be thought to be ambiguous, thus requiring any conclusions to be drawn to be treated with caution. As already noted the theoretical connection between this variable and the level of income is weaker than that for the index of earnings. Against this, the furthest one may go is to state that there must nevertheless be a presumption that the variable is related to income, since there is no significant evidence to the contrary, whilst the export variable is itself a function of the index of earnings variable.

It follows, therefore, that the e xplanation for the negative coefficient must be sought in the connection between the level of income and the size of the tertiary sector, and the various hypotheses tested in this respect are represented by equations (9.3) and (9.8) to (9.10). Again, however, the conclusions are disappointing. There is no evidence that productivity changes come between the level of income and the increase in the size of the tertiary sector (equation (9.3)), neither does it appear that a rise in income leads to a significant shift in expenditure towards housing (equations (9.9) and (9.10)). Moreover, the hypothesis that an increase in income will lead to a greater demand for imported services does not appear to be supported by the behaviour of the centres towards which this demand might be expected to be directed (equation (9.8)).

Hence, in conclusion to the negative coefficient of the income variables, there would appear to be no explanation capable of specification at the town level which can explain the negative sign whilst still remaining within the framework as laid down by traditional theory.

The remaining two aims of the last chapter, namely to develop new hypotheses and to refine the existing ones, are more successful. The index of commuting (equation (9.6)) shows that the geographical location of a town influences the size of the tertiary sector, not only through its relationship with larger towns (large town variable), but also through the extent to which it may draw upon a labour pool outwith its own boundaries. Nevertheless, perhaps the most interesting result is that obtained with the new hypothesis that towns possess tendencies in varying strengths towards being service centres. A large percentage of the variance (74%) is explained by the size of the tertiary sector 1931, and the suggestion that this measures the above influence is borne out by the behaviour of the hotel accommodation variable. Thus although the variables found to be significant in Chapter 7 explained 62% of the variance, this still leaves untouched the historical element, and it would appear that the hypothesis as embodied in equations (9.4) and (9.5) provides a satisfactory measure of the influence.

INTEGRATION OF THE RESULTS

It is now possible to consider the conclusions arrived at separately for the original model and for the extensions to it with a view to determining the best overall fit for the model. The results obtained for the three versions of the original model showed that the large towns, index of earnings and export sector variables represented the only significant relationships; whilst the preliminary investigations in the above section suggested that these should be augmented by the hypotheses represented by the indices of concentration and commuting, together with the 'historical' element as reflected in the size of the tertiary sector 1931 and the index of hotel accommodation.

The versions to be tested here are the two forms of the equilibrium model together with the disequilibrium model. In addition, the final modification suggested in Chapter 8, namely that the dependent variable should be specified in absolute rather than percentage terms, will be introduced and evaluated in this section.

Equilibrium Model

It will be recalled from Chapter 7 that the final form of the first equilibrium model was specified using the large town, index of earnings and export sector variables as

$$S_t = 10529 - \underset{(0.003)}{0.013}\, L_t - \underset{(2228)}{11581}\, T_t - \underset{(0.06)}{0.27}\, E_t \qquad (7.5)$$

Equation (7.5) had an $\bar{R}^2$ of 0.62. If this model is now extended to include the hypothesis that the size of the tertiary sector is a function of the degree of concentration in the older industries, then the variance explained increases to 66%. Thus -

$$S_t = 10131 - \underset{(0.003)}{0.01}\, L_t - \underset{(2416)}{8804}\, T_t - \underset{(0.06)}{0.28}\, E_t - \underset{(664)}{1372}\, C_t \qquad (9.23)$$

The main effect is to cause the coefficient upon the index of earnings to fall, but all the coefficients are still significant at the 95% confidence level. Thus the inclusion of the index of concentration combines with the other independent variables to explain a further 10.5% of the remaining variance.

If to equation (9.23) is now added the index of commuting, the variance explained rises by 1% to 67%, i.e. the inclusion of this variable explains a further 3% of the variance remaining in equation (9.23). The model may then be specified as -

$$S_t = 12409 - \underset{(0.003)}{0.009}\,L_t - \underset{(2399)}{7809}\,T_t - \underset{(0.06)}{0.27}\,E_t - \underset{(645)}{1402}\,C_t - \underset{(28.2)}{56.5}\,A_t \quad (9.24)$$

As might be expected from the behaviour of the $\overline{R}^2$, the coefficient upon the commuting variable is just significant at the 95% confidence level. The effect upon the other coefficients is to cause the coefficient on the large town variable to fall considerably, whilst its standard error remains the same. However it is still significant at the 95% level. In addition, of some interest is the constancy of the export sector coefficient in both equations (9.23) and (9.24).

The net effect of introducing the commuting variable is to improve marginally the fit of the model, whilst still retaining all the other variables as significant. However, as a result of collinearity between the commuting and the other independent variables, the addition of the former only explains 3% of the remaining variance compared to the 17% obtained in the single regression (equation (9.16)).

If the 'historical' element is considered as explaining the residue, then the size of the tertiary sector 1931 and the index of hotel accommodation should be added in last to the model. With the size of the tertiary sector 1931 the $\overline{R}^2$ rises from 0.67 to 0.82.[2] The model can then be specified as -

$$S_t = 6827 - \underset{(0.002)}{0.003}L_t - \underset{(1788)}{5975}T_t - \underset{(0.05)}{0.11}E_t - \underset{(506)}{269}C_t - \underset{(21.0)}{35.4}A_t + \underset{(0.08)}{0.55}S_{1931} \quad (9.25)$$

From an inspection of the coefficients and standard errors it can be seen that the inclusion of S_{1931} has a marked effect upon the remaining variables and the coefficients of the large town, concentration and commuting variables become insignificant at the 95% level. Thus in the presence of this variable (which by itself explained 74% of the variance), only the income variables remain significant.

With the addition of the alternative measure of the 'historical' element, the index of hotel accommodation, the variance explained rises 1% to 83% and the equation then becomes -

$$S_t = 6575 - \underset{(0.003)}{0.003}\,L_t - \underset{(1878)}{4862}\,T_t - \underset{(0.05)}{0.1}\,E_t - 159\,C_t - \underset{(20.7)}{37.9}\,A_t$$

$$+ \underset{(0.08)}{0.54}\,S_{1931} + \underset{(12.2)}{24.5}\,H_t \quad (9.26)$$

The coefficient on the index of hotel accommodation is just significant at the 95% level, whilst the significance of the other variables remains unchanged from equation (9.25).

If the 'historical' element is no longer treated as a residue and its importance recognised, then the variables should be added to the original model in the reverse order. With the inclusion of S_{1931} to

---

2. As already noted when this variable is introduced the sample has to be restricted and this might be thought to introduce possible error. However, when equation (9.24) is run again on this sample an $\overline{R}^2$ of 0.67 is again attained.

equation (7.5), the $\overline{R}^2$ rises from 0.62 to 0.82, i.e. the variance explained with only S_{1931} is now as high as that for equation (9.27) above. The model is then -

$$S_t = \underset{}{5135} - \underset{(0.003)}{0.003}\, L_t - \underset{(1687)}{6736}\, T_t - \underset{(0.05)}{0.10}\, E_t - \underset{(0.08)}{0.59}\, S_{1931} \qquad (9.27)$$

In this version, the inclusion of the size of the tertiary sector 1931 renders the large town variable of the original model insignificant. With the further inclusion of the alternative 'historical' element, the $\overline{R}^2$ rises to 0.83 thus -

$$S_t = 4834 - \underset{(0.002)}{0.003}\, L_t - \underset{(1813)}{5608}\, T_t - \underset{(0.05)}{0.10}\, E_t + \underset{(0.08)}{0.57}\, S_{1931} + \underset{(11.3)}{22.9}\, H_t \qquad (9.28)$$

In this instance the coefficient on the index of hotel accommodation is significant, but the coefficient of the large town variable remains insignificant at the 95% level.

With the inclusion of the remaining two hypotheses, as would be expected from equations (9.25) and (9.26), their coefficients are insignificant. In equation (9.29) below, the coefficient of the concentration index is insignificant and the $\overline{R}^2$ remains at 0.83.

$$S_t = 4876 - \underset{(0.002)}{0.003}\, L_t - \underset{(1888)}{5518}\, T_t - \underset{(0.05)}{0.10}\, E_t + \underset{(0.08)}{0.57}\, S_{1931} + \underset{(11.0)}{22.5}\, H_t - \underset{(51.2)}{100.2}\, C_t \qquad (9.29)$$

The same conclusion can be arrived at with the addition of the index of commuting, for the coefficients of the large town, concentration and commuting variables are all insignificant and the $\overline{R}^2$ is again 0.83 (i.e. equation (9.26) above).

In summary of the above analysis, it is apparent that the highly significant relationship between the size of the tertiary sector and the 'historical' element, principally in the form of S_{1931}, is sufficient to overwhelm the weaker, but hitherto significant, relationships between the large town, concentration and commuting indices and the dependent variable. It would appear that due to collinearity between these variables and the size of the tertiary sector 1931, their original relationships are too weak to remain significant in the presence of the latter.

The results serve to raise a dilemma in their interpretation. On the one hand a strict reading of the coefficients of the large towns, concentration and commuting variables would suggest that in the presence of the size of the tertiary sector 1931 they should be dropped from the analysis. Yet to follow this step might rob the analysis of certain hypotheses which, although not so in this case, could well contain significant explanatory powers for future situations. To put the point another way, in the context of predicting future industrial structures, to conclude that 45% of the remaining variance can be explained in terms of what the industrial structure was in the past may be thought to be unsatisfactory when there is a certain amount of evidence suggesting that alternative, more specific, hypotheses might explain part of this variance. Thus to accept this variable at the expense of the others might prove to be acceptable in a situation in which the industrial structure responded to natural changes in these variables, but, in a situation in which the initial changes are imposed from outside, to consider the future industrial structure in terms of what it was in the past would clearly be unacceptable (see for instance the original problem posed by the Central Borders in the Introduction).

Before a decision can be made at this point it is therefore necessary that the relationships should be considered in detail. In Chapter 7 it was shown that the two most important modifications lay in the division of the sample between north and south and the use of log transformation.

From Appendix B Table B21 it appears that the effect of the size of the tertiary sector 1931 in the individual north and south samples is broadly similar to the full sample (whether that variable is added in first or last).

For the south the final equation obtained was -

$$S_t = 4193 - \underset{(0.005)}{0.008}\, L_t - \underset{(2826)}{5335}\, T_t - \underset{(0.10)}{0.009}\, E_t - \underset{(1021)}{1075}\, C_t + \underset{(4.2)}{2.1}\, A_t + \underset{(0.16)}{0.60}\, S_{1931} + \underset{(23.9)}{14.4}\, H_t \quad (9.30)$$

The overall variance explained (0.80) is lower than for the full sample (equation (9.26)) but the effect of the size of the tertiary sector 1931 upon the significance of the other variables is even greater. Thus, upon the inclusion of this variable, in addition to the large town, concentration and commuting variables, the export and index of hotel accommodation also become insignificant at the 95% confidence level (caused by a fall in their coefficients).

The pattern in the north is slightly different as can be seen from the final model below.

$$S_t = 5275 - \underset{(0.005)}{0.007}\, L_t - \underset{(2173)}{5276}\, T_t - \underset{(0.10)}{0.009}\, E_t - \underset{(1021)}{1068}\, C_t + \underset{(2.4)}{5.6}\, A_t + \underset{(0.18)}{0.58}\, S_{1931} + \underset{(22.7)}{14.8}\, H_t \quad (9.31)$$

The overall variance explained increases to 84%[3] but the main point to note here is that, whilst the export variable is still insignificant, the index of commuting is not and that the latter explains 14% of the remaining variance. However, it would be difficult to place much reliance upon this result since the variable only yielded an $\bar{R}^2$ of 0.04 by itself, suggesting the presence of positive co-variance between this and the other variables. Furthermore, in the absence of a similar behaviour with the other samples, it must be concluded that this relationship with the other independent variables is peculiar to the north sample.

Thus these results fail to provide any evidence for an increased significance with the north/south samples either in the commuting or concentration indices or in the large town variable and the relationships are not substantially different from those of the full sample.

With the second modification suggested in Chapter 7, namely the use of log transformations, the results are similar to those obtained in that chapter for the original model. Whether the log transformation of the independent variables is tried against the dependent variable in its normal or its log form, the results are always worse than those obtained with the normal version of the independent and dependent variables (Appendix B Table B22). Moreover, the pattern of the other variables vis-à-vis the size of the tertiary sector 1931 is not altered in the sense that, in the presence of this variable, the coefficients of the large town, concentration and commuting indices become insignificant.

3. For this particular sample there was multicollinearity between the size of the tertiary sector 1931 and the index of concentration (zero-order correlation coefficient of 0.63). However, it is not expected that this is sufficient to affect the conclusions to be drawn.

So, despite these two modifications it would appear that the conclusions reached above that the large town, index of concentration and index of commuting variables should be deleted were correct and accordingly the model should be rewritten as -

$$S_t = F\left\{ T_t,\ E_t,\ S_{(t-1)},\ H_t \right\} \qquad (9.32)$$

However, when the model was actually tested in this form, equation (9.33) was obtained -

$$S_t = 4838 - \underset{(1826)}{5619}\,T_t - \underset{(0.06)}{0.10}\,E_t + \underset{(0.08)}{0.57}\,S_{1931} + \underset{(16.3)}{23.9}\,H_t \qquad (9.33)$$

The $\overline{R}^2$ was lower than that yielded by equation (9.26) and the coefficients upon the export and hotel accommodation variables became insignificant at the 95% level. Hence, as a result of collinearity between the independent variables, the model must be specified in terms of more variables than would be suggested by the behaviour of the $\overline{R}^2$ alone. Various combinations of equation (9.32) and the variables found to be insignificant were tried in order to find the best overall fit whilst retaining as few insignificant variables as possible. The version finally selected was equation (9.32) together with the commuting variable (A_t) -

$$S_t = 6326 - \underset{(1807)}{5028}\,T_t - \underset{(0.048)}{0.09}\,E_t + \underset{(0.07)}{0.59}\,S_{1931} - \underset{(20.3)}{40.8}\,A_t + \underset{(14.4)}{25.8}\,H_t \qquad (9.34)$$

Equation (9.34) has one variable (index of hotel accommodation) insignificant at the 95% confidence level (but significant at the 90% level) and explains 82% of the variance. It is this equation which may be taken as the final form of the first equilibrium version of the model.

To conclude the position with regard to the first form of the equilibrium model, it would appear that the results are in accordance with the suggestion that the percentage size of the tertiary sector is dependent negatively upon both the degree of concentration of the industrial structure in those industries which were important at the turn of the century and the percentage that the total labour force bears to the population. This is true whether the full sample, the north/south sample or log transformations are taken. Nevertheless the 'historical' element, as represented by the size of the tertiary sector 1931 and the index of hotel accommodation, appears to be so powerful that its inclusion not only renders the above two hypotheses insignificant, but the geographical variable as well. Furthermore, investigations based on the north/south sample together with the use of log transformations only serve to reinforce this result.

However, as a result of collinearity between the variables, the model was found to be best specified in a form which includes the index of commuting and, therefore, the only variables to be excluded as a result of the integration of the original and new hypotheses are those of the large town and index of concentration.

To turn attention to the first differences form of the equilibrium model, this requires using the 1951 data and it will be recalled from Chapter 6 that the statistics on industrial structure for this year were based on the 1948 standard industrial classification. Here the main effect of this is that the index of concentration (based as it is upon certain specified industries) could not be constructed for 1951 without radically altering its underlying assumptions. In order to preserve a consistency of approach with these additional variables, the levels of the indices of commuting and hotel accommodation were adopted. Moreover this decision was reinforced by the static nature of the index of hotel accommodation, together with the little change experienced in the commuting index.

The main conclusion to the first differences version of the original model (Chapter 7) was that only the change in the index of earnings was found to be significant. When the hypotheses extending this model are added in a manner similar to that adopted with the first version (above) the results of Appendix B Table B23 are obtained.

When the index of concentration is regressed against the change in the size of the tertiary sector, an $\bar{R}^2$ of 0.08 is obtained, together with a positive and significant coefficient viz. -

$$\Delta S_t = 38.5 + \underset{(281)}{684}\, C_t \tag{9.35}$$

However, for the index of commuting, the relationship was found to be insignificant viz. -

$$\Delta S_t = \underset{(13.4)}{20.9}\, A_t - 770 \tag{9.36}$$

As for the 'historical' element the size of the tertiary sector 1931 was both significant and gave the best overall fit ($\bar{R}^2 = 0.12$).

$$\Delta S_t = 735 - \underset{(0.04)}{0.11}\, S_{1931} \tag{9.37}$$

It should be noted that equation (9.37) carries a negative coefficient which is in contrast with the first equilibrium models, i.e. in this instance changes in the tertiary sector appear to be negatively related to their size in 1931.

On the other hand, the index of hotel accommodation proved to give an insignificant relationship, thus -

$$\Delta S_t = 228 - \underset{(8.7)}{10.4}\,H_t \qquad (9.38)$$

though as with equation (9.37) the coefficient is negative.

From Appendix B Table B23 it can be seen that the integration of these hypotheses with the original first differences model is unsatisfactory. The original version of this model was based upon the change in the index of earnings variable, i.e.

$$\Delta S_t = \underset{(943)}{2123}\,\Delta T_t - 6.51 \qquad (9.39)$$

and this equation was associated with an $\bar{R}^2$ of 0.07. However, the inclusion of the above hypotheses fails to produce a specification superior to that of equation (9.37). By way of illustration the model containing all these variables gave an $\bar{R}^2$ of 0.10 i.e.

$$\Delta S_t = 198 + \underset{(956)}{1958}\,\Delta T_t - \underset{(0.003)}{0.008}\,S_{1931} + \underset{(376)}{385}\,C_t + \underset{(0.22)}{0.01}\,A_t - \underset{(0.79)}{0.05}\,H_t \qquad (9.40)$$

Thus equation (9.40) fails to provide a better fit than equation (9.37), nor does the significance of the additional variables really justify the selection of this form of the model. Consequently equation (9.37) will be taken as the final specification of the first differences version of the model, since the introduction of new variables into the original model (equation (9.39)) fails to improve upon equation (9.37).

When the north/south samples and log transformations are tried, equation (9.37) is not improved upon (Appendix B Tables B24 and B25), so again the final specification should be in terms of equation (9.37).

One of the main reasons for pursuing the first differences version of the model was to assess the extent to which the first version of the equilibrium model was stable. Had the latter model been stable then any changes in the independent variables would have been associated with a change in the dependent variable. However, although the failure of the first differences equation is certainly consistent with instability in this sense, it does not necessarily follow that equation (9.34) is unstable, for all the above model (equation (9.37)) shows is that, for 1951 and 1961, the relationships were not identical. As such this difference could also be due to the 1961 version of equation (9.34) being in disequilibrium. However, now that the first version of the equilibrium model has been fully specified (equation (9.34)), it is possible to gain additional evidence by re-running the model using 1951 data and comparing the coefficients so obtained.

When this is performed the following equation results.

$$S_{1951} = \underset{(2876)}{51971}T_{1951} - \underset{(0.38)}{0.46}E_{1951} + \underset{(0.38)}{0.93}S_{1931} - \underset{(7.2)}{8.8}A_t + \underset{(2.8)}{4.6}H_t - 3196 \quad (9.41)$$

and this may be compared with the specification arrived at using the 1961 data, namely equation (9.34) -

$$S_{1961} = 6326 - \underset{(1807)}{5028}T_t - \underset{(0.048)}{0.095}E_t + \underset{(0.07)}{0.59}S_{1931} - \underset{(20.3)}{40.8}A_t + \underset{(14.4)}{25.8}H_t \quad (9.34)$$

From a comparison of equations (9.41) and (9.34) it will be observed that the coefficient of the index of earnings variable is significant but positive in equation (9.41), whereas in equation (9.34) it is significant but negative. Also the coefficients upon the export and index of commuting variables are insignificant at the 95% level in equation (9.41) but significant in (9.34). Finally, all the coefficients in equation (9.41) are significantly different from those in equation (9.34). Hence the difference in behaviour of the coefficients for 1951 and 1961 in the first equilibrium model lend considerable force to the argument that equation (9.34) is unstable. Unfortunately, however, it is impossible to be completely certain about this since part of the difference (at least as far as the index of earnings is concerned) may be due to the discrepancy between the 1948 and 1958 standard industrial classifications. Nevertheless, given the extent of the failure to specify the first differences model, as well as the behaviour of the coefficients, it can at least be stated that the stability of equation (9.34) is far from established.

Disequilibrium Model

In this section the new hypotheses may now be integrated with the disequilibrium version of the model. Since the disequilibrium model takes the same dependent variable as the first differences version of the equilibrium model, the remarks made about the 1951 data apply here as well and, as such, the model could only be tested using 1961 data (rather than using both years as did the original model in Chapter 7).

where 1951 not used due to changes classif 1948-1958

It will be recalled from Chapter 7 that no firm specification could be reached with the original model due to the lack of any relationship. This latter point was also reflected in the extremely low variance (1%) explained by that model. If the three new hypotheses are added to the model in the same manner as with the equilibrium models, the results obtained are those given in Appendix B Table B26. Since the dependent variable is the same as for the first differences version the single regression equations are the same as equations (9.35) to (9.38) above. When the index of concentration is added to the model based upon the large town, index of earnings and export sector variables, the $\overline{R}^2$ rises from 0.01 to 0.04, and the equation becomes

$$\Delta S_t = \underset{(0.002)}{0.0008}\, L_t + \underset{(1192)}{978}\, T_t - \underset{(0.03)}{0.007}\, E_t + \underset{(327)}{520}\, C_t - 145 \qquad (9.42)$$

From (9.42) it follows that all the coefficients are insignificant and that as such the positive coefficients upon the large town and index of earnings variables are not capable of interpretation. If the index of commuting is added to equation (9.42) all the coefficients are insignificant at the 95% level and the $\overline{R}^2$ remains at 0.04.

$$\Delta S_t = \underset{(0.002)}{0.0005}\, L_t + \underset{(1219)}{736}\, T_t - \underset{(0.03)}{0.009}\, E_t + \underset{(328)}{527}\, C_t + \underset{(14.4)}{13.7}\, A_t \qquad (9.43)$$

More surprising, perhaps, is the insignificance of the size of the tertiary sector 1931 when it is added into equation (9.43) since, by itself, it gave an $\overline{R}^2$ of 0.12 (see equation (9.37)).

$$\Delta S_t = 3906 - \underset{(0.002)}{0.0007}\, L_t + \underset{(1204)}{378}\, T_t - \underset{(0.03)}{0.04}\, E_t + \underset{(34.1)}{30.6}\, C_t + \underset{(14.1)}{9.6}\, A_t - \underset{(0.06)}{0.11}\, S_{1931} \qquad (9.44)$$

Moreover, the addition of the index of hotel accommodation fails to improve upon equation (9.44), thus -

$$\Delta S_t = 379 - \underset{(0.001)}{0.007}\, L_t + \underset{(1301)}{428}\, T_t - \underset{(0.03)}{0.04}\, E_t + \underset{(34.7)}{31.1}\, C_t + \underset{(14.3)}{9.5}\, A_t - \underset{(0.06)}{0.10}\, S_{1931} + \underset{(10.1)}{1.1}\, H_t \qquad (9.45)$$

It might be expected that, since the size of the tertiary sector 1931 gave the best fit, its addition into the original three variables might yield a significant result. This, however, was found not to be the case.

$$\Delta S_t = 935 - \underset{(0.002)}{0.0006}\, L_t + \underset{(1115)}{840}\, T_t - \underset{(0.03)}{0.04}\, E_t - \underset{(0.07)}{0.13}\, S_{1931} \qquad (9.46)$$

If the remaining variables are added into equation (9.46) in the reverse order to (9.45) no significant coefficients are obtained. Moreover, the lack of any significant relationship is confirmed when the north/south samples and log transformations are tried (Appendix B Tables B27 and B28).

It would, therefore, appear that no combination of variables improves upon the size of the tertiary sector 1931 as the sole independent variable and that the model should accordingly be specified as -

$$\Delta S_t = 735 - \underset{(0.04)}{0.11}\, S_{1931} \qquad (9.37)$$

Thus the new hypotheses developed in the last chapter, whilst giving a better fit, still do not give a satisfactory explanation of the change in the size of the tertiary sector. In the light of this more or less persistent failure of the disequilibrium model to yield any significant relationships either one of two conclusions may be drawn. First the original model may indeed be in equilibrium. If this were so then the percentage size of the tertiary sector as measured in both 1951 and 1961 would be fully adjusted to the independent variables, so that any change in the tertiary sector could not be as a response to the levels of these variables in either 1951 or 1961. Alternatively it may be hypothesised that the system is in disequilibrium and that it is therefore responding to the present levels of the independent variables, but that the process is one of partial adjustment. In this case the levels of the independent variables in 1961 would not have had sufficient time to induce a full response in the size of the tertiary sector, so that the changes now being measured would be part at least attributable to the value of the dependent variable at the outset of the period.

So instead of the model being -

$$\Delta S_t = F(Z) \tag{9.47}$$

where Z = the equilibrium model (equation (9.34)), it should now read

$$\Delta S_t = F(Z) + \phi S_{(t-1)} \tag{9.48}$$

However when the latter version is tested the conclusions are ambiguous. The coefficients of the equilibrium model (Z)(in equation (9.48)) are still insignificant, but that for the size of the tertiary sector at the beginning of the period is not and the $\bar{R}^2$ is increased to 0.21. So, although the change in the tertiary sector does not appear to be in response even partly to the levels of the independent variables in 1961, there is, nevertheless, some evidence to suggest the presence of a partial adjustment process.

From this it may be concluded that it is the second conclusion (above) that should be adopted and the disequilibrium model may accordingly be fully specified as[4] -

$$\Delta S_t = F(Z) - \underset{(0.09)}{0.213}\, S_{1951} \tag{9.49}$$

Absolute size of the tertiary sector

It will be recalled from Chapter 8 that one of the modifications suggested there was to define the dependent variable in terms of the absolute employment in the tertiary sector rather than as the percentage of the total labour force. In this case it is again possible to identify three versions of the model, viz. the two equilibrium models along

4. In this equation, the full specification of Z would be:

$$Z = 1532 - \underset{(189)}{179}\, T_t - \underset{(296)}{540}\, E_t + \underset{(0.04)}{0.038}\, S_{(t-1)} + \underset{(5.6)}{6.8}\, H_t$$

with the disequilibrium form. Hence, combining both the original and the new hypotheses previously found to be significant, the models may be specified as, using previous notation -

Equilibrium form:

$$S_{at} = F\left[P_t, L_t, E_t, T_t, S_{(t-1)}, C_t, A_t, H_t\right] \quad (9.50)$$

or
$$\Delta S_{at} = F\left[\Delta P_t, \Delta L_t, \Delta E_t, \Delta T_t, S_{(t-1)}, C_t, A_t, H_t\right] \quad (9.51)$$

Disequilibrium form:

$$\Delta S_{at} = F\left[P_t, L_t, E_t, T_t, S_{(t-1)}, C_t, A_t, H_t\right] \quad (9.52)$$

where S_{at} = the absolute size of the tertiary sector.

Equilibrium Model. The results of all the regressions for the first version of the equilibrium model are set out in Appendix B Tables B29 and B30 and the most important ones are summarised in Table (9.1) below.

Table 9.1

Results of single regression analysis with size of tertiary sector in absolute terms

INDEPENDENT VARIABLE	CONSTANT	COEFFICIENT	$\bar{R}^2$
Population (log)	975	0.09 (0.02)	0.70
Large town	41.325	-0.25 (0.09)	0.09
Index of earnings	40.224	-26078 (68793)	-0.01
Export Sector	127.6	0.01 (0.03)	-0.01
Size of tertiary sector 1931	1432	0.84 (0.07)	0.03
Index of Concentration	34652	- 3776 (19826)	-0.02
Index of commuting	21849	2.56 (9.7)	-0.02
Index of hotel accommodation	33794	- 20.6 (592)	-0.01
Labour force	10391	0.35 (0.04)	0.58

From Appendix B Table B29, with population (log), the full model yielded an $\overline{R}^2$ of 0.77, but despite this good overall fit the model may be thought to be unsatisfactory. This is due to the fact that, with the addition of the variables to the model specified in terms of population alone, only 20% of the remaining variance is explained, i.e. the $\overline{R}^2$ rises from 0.70 to 0.77. When it is realised that a large amount of the employment in the tertiary sector must of necessity be attributable to the town's population, the explanatory power of the model is seen to be rather disappointing.

In this connection it might be argued that the absolute size of the tertiary sector is more likely to be sensitive to changes in the size of the labour force rather than to the town's population. When this change is effected the path by which the final $\overline{R}^2$ is obtained is far more satisfactory as the equations below illustrate. Thus with the labour force alone, the $\overline{R}^2$ obtained is 0.58 and the equation may be specified as -

$$S_{at} = 10391 + \underset{(0.11)}{0.35}\, F_t \tag{9.53}$$

whilst the full model yields an $\overline{R}^2$ of 0.72.

$$S_{at} = 45365 + \underset{(0.04)}{0.39}\, F_t - \underset{(0.002)}{0.003}\, L_t - \underset{(48694)}{37032}\, T_t - \underset{(1.3)}{0.20}\, E_t + \underset{(1162)}{614}\, C_t$$

$$+ \underset{(2.1)}{5.3}\, S_{1931} - \underset{(587)}{1097}\, A_t - \underset{(401)}{1207}\, H_t \tag{9.54}$$

Although equation (9.54) gives a lower fit (0.72) than does the comparable model using population as an independent variable (0.77), it may nevertheless be preferred since the remaining variables raise the $\overline{R}^2$ from 0.58 to 0.72 i.e. they explain 36% of the remaining variance.

A detailed assessment of the significance of the variables in equation (9.54) has not been given since, for the first time in this analysis, the log transformations gave a better fit (Appendix B Table B31). The best relationship was found with the semi-log relationship, thus -

$$S_{at} = \underset{(1740)}{36455} \log F_t - 388960 \qquad (9.55)$$

For equation (9.55) the coefficient is significant and the $\overline{R}^2$ obtained is 0.89. However, with the subsequent inclusion of the remaining variables, the coefficient on log F_t is not significantly altered. In addition all the coefficients of these variables are insignificant at the 95% level. It would appear that the relationship between the labour force (log) and the dependent variable is so strong that it renders insignificant the weaker relationships based upon the other independent variables.

If the double log relationship is now investigated, the following set of equations is obtained. With the labour force alone, the coefficient is significant and the $\overline{R}^2$ is 0.70.

$$\log S_{at} = 1.6 + \underset{(0.07)}{0.83} \log F_t \qquad (9.56)$$

With the inclusion of the large town variable, the fit is unchanged,

$$\log S_{at} = 1.5 + \underset{(0.07)}{0.81} \log F_t + \underset{(0.02)}{0.001} \log L_t \qquad (9.57)$$

and the coefficient of the large town variable is insignificant. When the index of earnings variable is taken into equation (9.57) the $\overline{R}^2$ again remains unchanged, but this time the additional coefficient is significant at the 95% level.

$$\log S_{at} = \underset{}{1.7} + \underset{(0.07)}{0.83} \log F_t + \underset{(0.02)}{0.001} Lt - \underset{(0.9)}{2.2} \log T_t \qquad (9.58)$$

Similarly with the inclusion of the remaining income variable, the fit is unchanged.

$$\log S_{at} = 1.8 + \underset{(0.07)}{0.83} \log F_t + \underset{(0.01)}{0.0009} \log L_t - \underset{(1.0)}{1.8} \log T_t - \underset{(0.00003)}{0.000006} E_t \qquad (9.59)$$

However, when the two income variables are included both their coefficients become insignificant. The explanation for this latter point is not clear, since from an inspection of the zero-order correlation coefficients there does not appear to be any multicollinearity between these two independent variables.

The introduction of the size of the tertiary sector 1931 to equation (9.59) produces a significant change, since the $\overline{R}^2$ rises to 0.75 and the coefficient of this variable is significant (whilst those of the large town and income variables remain insignificant at the 95% level). Thus -

$$\log S_{at} = 1.9 + \underset{(0.07)}{0.80} \log F_t + \underset{(0.01)}{0.0009} \log L_t - \underset{(1.0)}{0.3} \log T_t + \underset{(0.00002)}{0.000008} E_t + \underset{(0.00004)}{0.0001} S_{1931} \qquad (9.60)$$

The inclusion of log S_{1931} thus causes the coefficient on log F_t to fall, but still to remain significant, whilst the former variable explains 13% of the variance remaining in equation (9.59). Finally, with the addition of the alternative measure of the 'historical' element, the index of hotel accommodation, the $\overline{R}^2$ is raised to 0.80, i.e. it explains 20% of the remaining variance in equation (9.60). The model then is -

$$\log S_{at} = 1.7 + \underset{(0.07)}{0.86} \log F_t + \underset{(0.01)}{0.0009} \log L_t - \underset{(0.9)}{0.09} \log T_t$$

$$+ \underset{(0.000002)}{0.000004} \log E_t + \underset{(0.00005)}{0.0001} \log S_{1931} - \underset{(0.01)}{0.03} H_t \quad (9.61)$$

The results obtained with equation (9.55) and equations (9.56) to (9.61) pose a dilemma in their interpretation. Equation (9.55) gives the best fit but the remaining variables explain none of the variance, whilst equation (9.61) gives a lower $\overline{R}^2$ (0.80) but contains more significant variables. Although one must not be too preoccupied in attaining the best fit, it was felt in this instance that since equation (9.55) explained 45% of the remaining variance in equation (9.61) the unsatisfactory behaviour of the remaining variables should be given less weight and a choice made in favour of equation (9.55).

From this it follows that the first version of the equilibrium model should be specified as in equation (9.55).

If attention may now be directed to the first differences version of the equilibrium model the results in Appendix B Table B32 are obtained. For the single regression analysis, these equations are reproduced in Table (9.2) below.

Table 9.2

Single regression analysis for first differences version based upon the absolute size of the tertiary sector

INDEPENDENT VARIABLE	CONSTANT	COEFFICIENT	$\bar{R}^2$
Change in labour force	827	0.28 (0.03)	0.57
Change in large town	1330	0.08 (0.13)	- 0.01
Change in index of earnings	2619	-13151 (3425)	-0.03
Change in export sector	1314	0.16 (0.47)	-0.02
Index of concentration	4120	3957 (4003)	-0.02
Index of commuting	8090	-185 (178)	-0.01
Size of tertiary sector 1931	7088	-0.38 (0.28)	-0.01
Index of hotel accommodation	4921	36.5 Ht (19.8)	-0.02

From Table 9.2. it can be seen that only the change in the size of the labour force provides any significant relationship. With this variable the coefficient is positive, illustrating that increases in the employment in the tertiary sector have been associated with increases in the size of the labour force.

When the remaining variables are considered alongside this change in the size of the labour force, the coefficients on these variables are always found to be insignificant; whilst that upon the change in the size of the labour force actually rises from 0.28 to 0.29 and still remains significant at the 95% confidence level. This point is illustrated in the specification obtained with all these variables taken together, thus -

$$\Delta S_{at} = 9713 - \underset{(0.05)}{0.29}\Delta F_t + \underset{(0.1)}{0.01}\Delta L_t - \underset{(11096)}{3776}\Delta T_t + \underset{(0.32)}{0.37}\Delta E_t$$

$$+ \underset{(3267)}{3884}\, C_t - \underset{(0.47)}{0.25}\, S_{1931} - \underset{(140)}{167}\, A_t - \underset{(100.5)}{36.5}\, H_t \qquad (9.62)$$

The main reason for testing the first differences version of the model is to discover the extent to which the relationship obtained with the first version of the equilibrium model is stable. It is, therefore, encouraging to note that the $\bar{R}^2$ associated with the labour force alone in the first equilibrium version (0.58) is very close to the $\bar{R}^2$ of 0.57 yielded by the first differences model. The presence of a stable relationship is further indicated when the coefficients for the labour force singly in the first version are compared for the years 1951 and 1961. These are, respectively, 0.347 and 0.339 and, with a standard error for 1961 of 0.039, these two coefficients are not significantly different, illustrating the fact that, over this period, the relationship has remained unaltered.

Unfortunately the same analysis cannot be extended to equation (9.~~52~~55) since some of the changes in the size of the labour force were negative, thus requiring the zero to be re-set and the coefficients thereby altered before log transformations could be used. Nevertheless, if the normal version is stable then it may be legitimately presumed that the log version is also stable, so to the extent that this presumption is correct, this step is not strictly necessary anyway.

Consequently, for the purposes here, the model may be specified as, from the results given in Table 9.2 -

$$\Delta S_{at} = 827 + \underset{(0.03)}{0.28}\,\Delta F_t \tag{9.63}$$

In conclusion to the equilibrium models based upon the absolute size of the tertiary sector, it would appear that the best relationship is that obtained with the log version of the labour force in the first equilibrium version (i.e. equation (9.55)), and that from the behaviour of the coefficients in 1951 and 1961, together with the first differences model, this is apparently stable over the period 1951 to 1961.

Disequilibrium Model

When equation (9.52) is evaluated the results obtained are set out in Appendix B Table B33 and the single regression equations are given in Table 9.3 below for the labour force, large town and income variables (the remaining hypotheses are the same as for Table 9.2 above).

Table 9.3

Results for single regression analysis in disequilibrium model based upon the absolute size of the tertiary sector

INDEPENDENT VARIABLE	CONSTANT	COEFFICIENT	$\bar{R}^2$
Labour force	-525	0.02 (0.02)	-0.02
Large town	5383	-0.005 (0.04)	-0.01
Export sector	14740	-1.98 (0.6)	0.14
Index of earnings	7800	-11406 (30758)	-0.01

This table shows that only the export sector variable is significant at the 95% confidence level. It should be noted that the coefficient is again negative, thus agreeing with the general inverse relationship between the income variables and the dependent variables.

As with the first differences version of the equilibrium model, the inclusion of the remaining variables fails to add to the relationship based upon the single significant variable. The full model is specified in equation (9.64) below, from which it appears that the coefficient upon the export sector variable has actually risen.

$$S_{at} = 32366 - \underset{(0.8)}{2.3}\,E_t - \underset{(0.04)}{0.005}\,L_t - \underset{(30655)}{11399}\,T_t - \underset{(8089)}{3021}\,C_t$$
$$- \underset{(1.4)}{1.16}\,S_{1931} + \underset{(0.02)}{0.01}\,F_t - \underset{(369)}{139.9}\,A_t - \underset{(243)}{111}\,H_t \qquad (9.64)$$

Thus it would appear that there is no collinearity between the remaining independent variables and that together they fail to improve upon the overall fit. Furthermore, this is a conclusion which is reinforced when log transformations are tried (Appendix B Table B34). In addition, unlike the version based upon the percentage change in the tertiary sector (equation (9.49)), there is no evidence for any adjustment process, since adding the size of the labour force 1951 fails to improve upon the fit of the model.

Accordingly, on the basis of these results, the disequilibrium model should be specified as:

$$S_{at} = 14740 - \underset{(0.06)}{1.98}\,E_t \qquad (9.65)$$

Thus, the size of the export sector apart, this section shows that not only is equation (9.50) stable but it also seems to be in equilibrium. It is for these reasons that the version based upon this model (i.e.equation (9.55)) shall be taken as the main conclusion for the absolute size of the tertiary sector.

Chapter Ten

CONCLUSIONS

Starting with the problem of forecasting industrial structures in general, the aim of this thesis was re-specified in terms of the size of the tertiary sector and the study was based upon a cross-section analysis of towns with populations in excess of 50,000 lying outwith the influence of London and the conurbations. The model was initially developed from theoretical considerations alone. It was then modified in the light of the results so obtained and the final model specified in the form which provided the best overall fit. To this end two equilibrium versions were selected, along with one disequilibrium form, and one in which the dependent variable was specified as the absolute numbers employed in the tertiary sector rather than as its size as a percentage of the total labour force. In all these versions the tertiary sector was defined as industrial orders 18 to 24. Finally, although the model was tried with a variety of samples, on the whole the best results for all four versions were obtained with the full sample.

The four versions were finally specified as follows:

(a) <u>Equilibrium version</u>

(a) $$S_t = 6326 - \underset{(1807)}{5028}\,T_t - \underset{(0.048)}{0.095}\,E_t + \underset{(0.07)}{0.59}\,S_{1931} - \underset{(20.3)}{40.8}\,A_t + \underset{(14.4)}{25.8}\,H_t \qquad \left[\bar{R}^2 = 0.82\right] \qquad (9.34)$$

(b) $$\Delta S_t = 735 - \underset{(0.04)}{0.11}\,S_{1931} \qquad \left[\bar{R}^2 = 0.12\right] \qquad (9.37)$$

× (b) <u>Disequilibrium version</u> $$\Delta S_t = F(Z)^{1} - \underset{(0.09)}{0.213}\,S_{1951} \qquad \left[\bar{R}^2 = 0.21\right] \qquad (9.49)$$

---

1. The full specification of Z is

$$Z = 1532 - \underset{(189)}{179}\,T_t - \underset{(296)}{540}\,E_t + \underset{(0.04)}{0.04}\,S_{(t-1)} + \underset{(5.6)}{6.8}\,H_t$$

× Better than 175t (using 1951 data) (i) diff 0.21 v 0.26 not significant
(ii) has more variables

(c) Absolute Employment version

$$S_{at} = \underset{(1740)}{36455} \log F_t - 388960 \qquad \left[\bar{R}^2 = 0.89\right] \qquad (9.55)$$

The conclusions to be drawn from each of these models may be considered in turn.

1) Equilibrium Models

To take the version based upon the percentage size of employment in the tertiary sector 1961 (equation (9.34)) first, the conclusions can be divided into those to be drawn from a) the preliminary investigations, b) the variables omitted in the final specification and c) the variables retained.

First, in the preliminary investigations it was found that, at the county level, the index of earnings and the non-domestic rateable value variables were satisfactory proxy variables for income. Furthermore, the tentative conclusion finally arrived at in connection with the export sector variable was that it too was connected with income. In addition, the preliminary investigations were concerned with determining the correct 'weights' to be attached to the population and geographical location variables. In the light of the results obtained it would appear that either the gravity formula provides no further sophistication or that it does but that this is masked by the lack of an overall relationship. However, in view of the $\bar{R}^2$ of 0.20 obtained with the large town variable, together with its insensitivity towards the weighting factor, it is suggested that the former is the correct interpretation.

To consider the variables omitted from either the theoretical model (equation (7.1)) or the new hypotheses found to be significant in the first section of Chapter 9, these are the following:

1) Population - the original hypothesis advanced was that towns would be organised into hierarchies and that the higher up a town was the larger would be its tertiary sector. In addition it was further postulated that the population of the town could be used to identify its position in the hierarchy. Although it was possible to construct a sample of 38 towns for which this variable was important, the overall fit of the model was still worse than that for the full sample. Moreover, the influence of the population variable was clearly shown to be inconsistent and to be capable of easily being outweighed by the presence of other variables.

2) Geographical variables - it was also suggested that the position of a town in the hierarchy would be influenced by its interaction with the towns nearby. Thus the presence of small towns would tend to increase the size of the tertiary sector and vice versa for the presence of nearby large towns. However it was only the latter which was found to be correct, but even its influence was rendered insignificant by the presence of the 'historical' element. From this it may be concluded that the geographical location of the town does have an influence upon the size of the tertiary sector, but that this influence is insignificant in the presence of other variables. It follows, therefore, that the variable does not improve upon the overall fit of the model and accordingly should be deleted.

From the conclusions to be drawn for both the population and geographical location variables it can be seen that the hypotheses put forward in Chapter 3 based upon the notion of an hierarchical system of towns may be correct in themselves, but when considered alongside other variables they fail to improve upon the overall fit to the model.

3) Concentration in older industries - a similar conclusion can be drawn from the behaviour of the variable measuring the extent to which a town's industrial structure is concentrated upon activities which were deemed to be important 50 to 60 years ago. The size of the tertiary sector varies inversely with this index alone, indicating that those towns which have this type of industrial structure do in fact have smaller tertiary sectors. Yet again when this variable is added into the model it fails to improve upon the overall fit and, for this reason, it is not included in the final specification.

To treat the results obtained with this variable with caution would appear to be justified by the behaviour of the change in population variables. It was argued that if towns did have such a high concentration in one or two older industries, then, for a number of reasons, towns would be associated with low population increases, but when this hypothesis was tested no evidence for this was found. Hence the alternative method of specifying the hypothesis underlying the index of concentration gave no results, thus casting doubt upon the viability of that hypothesis.

Finally, to turn attention to those hypotheses included in the final specification of the model, it can be seen from equation (10.1) that these consist of the income and 'historical' variables, along with the index of commuting. Each may be considered in turn.

a) Income variables - these consist of both the index of earnings in the secondary sector and the size of the exporting sector. For these the two conclusions to be drawn are that both are significant and both carry a negative coefficient. Moreover the latter would appear to be invariant with the samples chosen, or with the degree of disaggregation adopted in the dependent variable. Neither would it appear to be due to any mis-specification of the variables.

The original hypotheses behind these two variables are:

a) that they are 'proxy' variables for income. For the index of earnings in the secondary sector this is a direct relationship, whilst for the export sector it is based upon the logic of the regional multiplier, viz. that the higher the exporting activity the greater will be the flow of income into the region.

b) that the higher the level of income the larger will be the tertiary sector in employment terms due to the higher income elasticities of tertiary products along with the smaller scope for productivity increases. However the negative coefficient upon both these variables disputes the combined effect of (a) and (b) and explanations for this have been sought in a failure of the relationship between the proxy variables and the level of income (hypothesis (a)) and/or a break in the relationship between the level of income and the size of the tertiary sector (hypothesis (b)).

It is only with the export sector that there is any real possibility of a failure in hypothesis (a) since it was concluded in Chapter 7 that the index of earnings was an acceptable proxy variable for income. The use of the export sector poses many statistical difficulties, the chief of which is the possibility that towns with large export sectors are also those which, for historical reasons, have large secondary sectors. However, as already noted, too much reliance cannot be placed upon the index of concentration (which was devised to test this hypothesis) and the corollary to this argument of a low population change has also been firmly rejected. Hence there is no real evidence to suggest a failure in hypothesis (a) above.

The search for a failure in hypothesis (b) is also unsatisfactory for no evidence can be found that productivity changes come between the level of income and the size of the tertiary sector, neither would it appear that, as suggested by the family budget surveys, an increase in income would be directed towards housing (which is outwith the tertiary sector as defined here). Moreover the hypothesis that an increase in income would lead to a disproportionate increase in the demand for imported services would appear to be unsubstantiated.

Thus there would appear to be no explanation for the negative coefficient upon the income variable that is capable of verification at the town level. Yet, since both variables are significant even in the presence of all the other hypotheses, and since the behaviour of these variables is so consistent, they must be included in the final specification of the model.

b) 'Historical' variables - the results for both the size of the tertiary sector 1931 and the index of hotel accommodation variables strongly reinforce the concept of the presence of an 'historical' element influencing the size of the tertiary sector. It would appear, therefore, that towns which have developed the reputation or characteristic of being service centres, as measured either by the size of the tertiary sector 1931 or,

more directly, by their provision of hotel accommodation, will continue to be influenced by this characteristic. As such it is reasonable to suggest that it is this factor which is responsible for the wide disparity between the sizes of the tertiary sector in otherwise similar towns.
c) Commuting variable - the negative coefficient of this variable indicates that the greater the extent to which a town relies upon its labour force, i.e. the smaller the degree of commuting, the lower will be the size of the tertiary sector. Hence it may be concluded that the commuters themselves provide an important element in the demand facing the tertiary sector.

In final summary to this version of the equilibrium model, it may be stated that, for the year 1961, 82% of the variation in the percentage size of the tertiary sector between towns of populations in excess of 50,000 lying outwith the influence of either London or conurbations can be explained. However, in so doing, the final model casts considerable doubt upon the traditional theoretical framework as developed in Chapters 2-5, for it shows that the concept of towns having certain functions and of being organised in an hierarchy, as location theory and its derivatives would suggest, plays little part in explaining the variance. More importantly, perhaps, is the finding that although the income variables are significant and are capable of explaining a considerable portion of the variance, they do so only in a manner contradictory to traditional theory.

2) First Differences Model

The final specification of the model, in terms of obtaining a good overall fit, is unsatisfactory and accordingly little may be concluded from the model per se. However, equation (9.37) is valuable in the sense that it serves to indicate that equation (9.34) may well be unstable, and this is a point which is further reinforced by the behaviour of the coefficients for equation (9.34) when it is run with the 1951 and 1961 data.

3) Disequilibrium Model

This version has proved to be the most intractable, and, as such, the conclusions to be drawn must of necessity be tentative in their nature. Perhaps the best way in which these may be presented is to summarise the evidence both for and against the presence of a disequilibrium situation. The evidence against this is, simply, the failure to adequately specify the disequilibrium model whether in terms of the percentage or the absolute size of the tertiary sector 1961. If equation (9.34) did specify an equilibrium position, then at each point of time the adjustment to past and present levels of the independent variables would be complete, so that no relationship should be observable when the levels of the independent variables are related to the change in the dependent variable over the period. Hence the failure here of the disequilibrium model might be put forward as evidence for such an equilibrium situation.

However, in opposition to this, the following should be noted:

a) The failure of the disequilibrium model is a necessary condition for the presence of an equilibrium position, but in the sense that this failure might be due to other factors it is not a sufficient condition.

The above logic implies that there is a complete response by the dependent variable, so that there should be no change in that variable which was not attributable to changes in the independent variables. Yet in the light of the considerable variation experienced by this dependent variable over the period 1951 to 1961 this is manifestly not the case. Although part of the change may be due to changes in the variables governing the still unexplained variance in equation (9.34), it would be more than fortunate if the whole of the change in the dependent variable was due to this cause.

b) The argument for an equilibrium position presupposes an instantaneous movement towards that position which, given the flexibility in output of the work force in the tertiary sector, is certainly unrealistic.

c) Although the disequilibrium model could not be specified satisfactorily, there is still some evidence for a disequilibrium situation. This can be partly found in the $\overline{R}^2$ of 0.24 achieved by the theoretical model using 1951 data, but is mainly due to the appearance in equation (9.49) of a partial adjustment process.

The net effect of these arguments is that it is unlikely that the relationship specified in equation (9.34) is in complete equilibrium, but unfortunately the failure of the disequilibrium model renders the precise nature of the disequilibrium process unidentifiable.

4) <u>Absolute size of the tertiary sector</u>

The above three models were repeated with the absolute, rather than the percentage, size of the tertiary sector as the dependent variable. Its best overall fit was found to be obtained with the first equilibrium version taking as the only independent variable the log transformation of the size of the total labour force (equation (9.55)). Moreover, from an inspection of the coefficients obtained in 1951 and 1961, together with the results yielded by the first differences model, it may be concluded that this model is stable. Finally, although as with the other models any conclusions drawn concerning the state of disequilibrium must be tentative, it would appear that there is less evidence to suggest that the relationship is not one of equilibrium.

What equation (9.55) clearly shows is that the total size of the labour force plays a dominating role in determining the absolute size of the tertiary sector and the apparent stability of the relationship may be taken as a further sign of its overall strength.

FINAL COMMENT

The conclusions so far have been presented in terms of the equilibrium and disequilibrium models. However, in any practical application of these results, the choice of which model to select is likely to be conditioned by the appropriateness or otherwise of each of the dependent variables to the particular problem at hand, and for this reason it may be worthwhile to consider the final position with regard to each of these variables.

If it is desired to predict the size of the tertiary sector at any point of time, then this may be achieved through considering either the percentage size or the absolute employment in the tertiary sector. Although the absolute employment model provides the best fit and appears to be both in equilibrium and stable, it may nevertheless be considered operationally weaker than the percentage size form since the latter encompasses a wider variety of influences. Against this, however, may be raised the evidence that this version, as specified in 1961, is in disequilibrium and may well be unstable.

On the other hand, if the analyst requires to measure the change in the size of the tertiary sector the position is less satisfactory. This change may be measured either in percentage or in absolute terms and both of these may be related to the levels of the independent variables at the beginning or the end of the period or to changes in these variables through the period. Yet despite this variety of alternative specifications it would appear that none furnish a viable practical alternative.

Finally the contribution offered by these models may be assessed in terms of the progress made in overcoming the initial problem posed in the Introduction, namely the inability of practical studies to predict the size of the tertiary sector other than in terms of falling within the range 40-60%. In the light of the above models it is now at least possible to identify some of the major factors affecting the size of the tertiary sector, together with the extent of their influence. Moreover since the basic unit of analysis, the town, is relatively small the size of a region's tertiary sector may be assessed more accurately through a detailed study of its individual components rather than, as in the past, through a general review of the whole regional economy. So, although firm forecasting per se may not always be possible due to the controversial nature of some of the assumptions relating to both the techniques employed and the data available, one may still be tempted to agree with J. S. Mill that "knowledge insufficient for prediction may be most valuable for guidance'. (Logic Book 6, Chapter IX)

BIBLIOGRAPHY

1. ALLEN, G.R. - 'The "Courbe des populations". A further analysis.' B.O.U.I.S. 16. 1954.

2. BAUER, P.T. and YAMEY, B.S. - 'Economic progress and occupational distribution.' E.J. 61. 1957.

3. BERRY, B.J. and GARRISON, W.L. - 'The functional basis of central place hierarchy.' E.G. 34. 1958.

4. BOGUE, D.J. - 'The structure of metropolitan community.' University of Michigan. 1949.

5. BRUSH, J.E. - 'The hierarchy of central places in south-western Wisconsin.' G.R. 43. 1953.

6. CARRUTHERS, W.J. - 'A classification of service centres in England and Wales.' G.J. 123. 1957.

7. CHRISTALLER - 'An hierarchy in economic location.' G.R. 23. 1938.

8. CLARK, C. - 'The condition of economic progress.' Macmillan & Co. London. 1940.

9. CLARK, C. - 'The location of industry and population.' T.P.R. 35. 1964.

10. CLARK, C. - 'Industrial location and economic potential.' L.B.R. 82. 1966.

11. CZAMANSKI, S.- 'Industrial location and urban growth. T.P.R.36. 1965.

12. DUNCAN, O.D. et al. - 'Metropolis and Region'. John Hopkins, Baltimore. 1960.

13. DUNNING, J.H. - 'Economic Planning and Town Expansion'. Workers' Educational Association. Southampton. 1963.

14. FISHER, A.G.B. - 'Capital and growth of knowledge.' E.J. 1933.

15. FLEMING, J.B. and GREEN, F.H.W. - 'Some relations between country and town in Scotland.' S.G.M. 68. 1952.

16. FOOTE, M.N. and HALL, P.K. - 'Social mobility and economic advancement.' A.F.R. 43. 1953.

17. FRIEDLY, F. - 'A note on the retail trade multiplier and residential mobility.' J.R.S. 1965.

18. FUCHS, V.R. - 'The growing importance of the service industries.' J.B. 38. 1965.

19. GREEN, F.H.W. - 'Community of interest areas - notes on the hierarchy of central places and their hinterlands.' E.G. 34. 1951.

20. GREENHUT, M. - 'When is the demand factor of location important?' L.E. 40. 1964.

21. HARRIS, C. - 'Functional classification of cities in the U.S.' G.R. 33. 1943.

22. HILDEBRAND, G.H. and MACE, A. - 'The employment multiplier in an expanding market 1940-47.' R.E.S. 1950.

23. HOOVER, E.M. - 'Location of economic activity.' McGraw-Hill. N.Y. 1948.

24. HOOVER, E.M. - 'The concept of a system of cities.' E.D.C.C.3. 1954.

25. HOUSING and LOCAL GOVERNMENT (Ministry of) - 'Administrative Areas Map.' 1966.

26. ISARD, W. - 'Location and space economy.' John Wiley. N.Y. 1956.

27. ISARD, W. - 'Methods of Regional Analysis.' John Wiley. N.Y. 1960.

28. ISARD, W., SCHOOLER, F. and VIETORISZ - 'Industrial complex analysis and regional development.' John Wiley. N.Y. 1959.

29. KUZNETS, S. - 'Gross national product since 1869.' N.Y. 1946.

30. LEONTIEF, W. and STROUT, A. - 'Multi-regional input-output analysis.' Paper presented to international conference on input-output techniques. Geneva. 1961.

31. LOASBY, B. - 'Making location policy work.' L.B.R. 83. 1967.

32. LÖSCH, A. - 'The economics of location.' Yale University Press. New Haven. 1954.

33. MACKAY, J.P.H. and CLYDE, J.J. - 'Armour on valuation for rating.' Green. Edinburgh. 1961.

34. MAJUMDAR, T. - 'The measurement of utility.' Macmillan. 1958.

35. MEYER, J.R. - 'Regional Economics - a survey.' Survey of Economic Theory II. Macmillan. 1965.

36. MOORE, F.T. and PETERSON, J.W. - 'Inter-Industry Model of Utah.' R.E.S. 1955.

37. MOSER, C.A. and SCOTT, W. - 'British Towns.' Oliver & Boyd. Edinburgh. 1961.

38. NEIDERCORN and KAIN - 'Econometric model of metropolitan development.' Regional Science Association Papers Vol.II. 1963.

39. NELSON, H.J. - 'A services classification of American cities.' E.G. 31. 1955.

40. NORTH, D.C. - 'Exports and regional growth - a reply.' J.P.E. 66. 1956.

41. PERLOFF, H.S. - 'Interrelations of state income and industrial structure.' R.E.S. 35. 1953.

42. PERLOFF, H.S. et al. - 'Regions, resources and economic growth.' John Hopkins. N.Y. 1960.

43. PREDÖHL - 'The theory of location in its relation to general economics.' J.P.E. 36. 1928.

44. REGAN, W.J. - 'Economic growth and services.' J.B. 36. 1963.

45. REYNOLDS, J. - 'Shopping in north-west.' T.P.R. 34. 1963.

46. ROTTENBERG, S. - 'Note on "economic progress and occupational distribution".' R.E.S. 35. 1953.

47. SCOTTISH COUNCIL - 'Inquiry into Scottish economy,1960-61.' Edin.1963.

48. SCOTTISH DEVELOPMENT DEPARTMENT - 'Lothians Regional Survey and Plan.' 1966.

49. SCOTTISH DEVELOPMENT DEPARTMENT - 'Central Borders.' 1968.

50. SIDDAL, W.A. - 'Wholesale-retail ratios as indices of urban centrality.' E.G. 31. 1961.

51. SMAILES, A.E. - 'The urban hierarchy in England and Wales.' G.29. 1944.

52. STEWART, C.T. - 'Size and spacing of cities.' G.R. 48. 1958.

53. STEWART, C.T. and WARNTZ, W. - 'Physics of population distribution.' J.R.S. II. 1958.

54. STIGLER, G.J. - 'Trends in employment in the service industries.' Princeton University Press. 1960.

55. STOPLER - 'Economics of location.' (Book Review) A.E.R. 43. 1965.

56. TAAFFE, E.J. - 'Urban hierarchy - air passenger definition.' E.G. 38. 1962.

57. THOMAS, M.D. - 'Export base and development stages theories of regional economic growth.' L.E. 40. 1957.

58. TIEBOUT, C.M. - 'Exports and regional economic growth.' J.P.E.66. 1956.

59. TROTMAN, R. and DICKENSON, D. - 'Scottish industrial estates.' S.J.P.E. 1961.

60. ULLMAN, E. - 'A theory of location of cities.' A.J.S. 46. 1941.

61. UZAWA, H. - 'On a two-sector model of e conomic growth.' Review Economic Studies. 29. 1961.

62. VINING, R. - 'A description of certain aspects of an economic system.' E.D.C.C. 3. 1954.

63. WEBER, A. - 'Theory of location of industries.' University of Chicago. 1929. English edition with introduction by C.J. FRIEDRICH.

64. WELSH OFFICE - 'A new town in mid-Wales.' H.M.S.O. 1965.

65. ZIPF, G.K. - 'Human behaviour and the principle of least effort.' Addison & Wesley Press. Cambridge, Mass. 1949.

KEY

A.E.R.	American Economic Review
A.F.R.	American Finance Review
A.J.S.	American Journal of Sociology
B.O.U.I.S.	Bulletin of Oxford University Institute of Statistics
E.D.C.C.	Economic Development and Cultural Change
E.G.	Economic Geography
E.J.	Economic Journal
G.	Geography
G.J.	Geographic Journal
G.R.	Geographic Review
J.B.	Journal of Business
J.P.E.	Journal of Political Economy
J.R.S.	Journal of Regional Science
L.B.R.	Lloyds Bank Review
L.E.	Land Economics
R.E.S.	Review of Economics and Statistics
S.G.M.	Scottish Geographical Magazine
S.J.P.E.	Scottish Journal of Political Economy
T.P.R.	Town Planning Review

APPENDIX A

THE DATA

TABLE AI

Data is presented in this table for the neighbourhood, small town and large town variables in that order. For each variable the data consists of the name of the place, its population and distance from the sample town.

ABERDEEN 185,390

Cults	3,910	3.5	-			-		

BARNSLEY 74,704

Barton	2,648	3.0	Dewsbury	53,490	10	Doncaster	86,322	15
Darfield	6,881	4.5	Hemsworth	14,500	8	Huddersfield	130,000	17
Dodworth	2,597	2.5	Holmefirth	18,620	15	Leeds	510,000	19
Houghton	1,008	5.0	Mexborough	16,680	16	Manchester	661,000	36
Royston	8,485	3.5	Pontefract	26,300	14	Sheffield	494,344	14
Hendley	1,425	1.5	Wakefield	61,628	10			
Wombwell	18,877	3.5						

BARROW 64,927

Dalton	10,316	4.0	Ulverston	10,310	9	-		

BATH 80,901

Bathford	1,281	3.0	Chipperham	18,510	13	Bristol	437,408	13
Bilton	3,652	5.0	Frome	11,700	15			
Bradford	5,760	5.0	Warminster	10,420	19			
Claverton	100	2.5						
Cold Ashton	242	4.5						
Englishcombe	351	2.5						
Freshford	594	3.5						
Coombe	2,330	2.0						
Farleigh	452	3.5						
Newton	242	3.0						
Pleasedown	3,238	5.0						
Stoke	464	2.0						
Wellow	408	3.5						
Westwood	771	4.5						

BEDFORD 63,334

Bromham	2,722	2.5	Huntingdon	11,480	21	Cambridge	95,527	29
Cardington	308	3.0	Kettering	38,840	24	Luton	131,000	19
Kempston	1,289	1.5	Leighton	12,280	20	Northampton	105,421	21
Ravensden	525	3.5	Bushden	17,490	13			
Renhold	739	3.5	Wellingborough	31,910	18			
Stagsden	400	4.5						
Stewartby	1,235	5.0						
Wittington	510	4.0						
Wilsthampstead	1,095	4.0						

BLACKBURN 106,242

Accrington	39,018	5.0	Burnley	80,559	11	Liverpool	747,000	39
Clayton	6,421	5.0	Chorley	31,060	10	Manchester	661,000	24
Darwen	14,142	4.5	Clitheroe	12,550	11	Preston	113,341	10
Harwood	10,718	5.0	Hastingdean	14,210	8			
Mellor	1,387	3.0	Rawtershall	12,550	11			
Os'whistle	3,906	3.5						
Rushton	5,433	3.5						
Salesbury	283	3.5						
Wilpshire	1,933	3.0						
Withnell	2,849	5.0						

BLACKPOOL 153,185

Cleverlys	20,468	4.5	Fleetwood	28,460	9	Liverpool	747,000	48
Lytham-St-Annes	36,189	4.5	Lancaster	48,400	26	Manchester	661,000	47
Poulton	6,009	3.5	Morecambe	40,500	17			
-			Preston	113,341	17			

BOURNEMOUTH 154,296

Christchurch	26,336	5.0	Dorchester	13,200	27	-		
Hamperston	6,534	5.0	Lymington	30,610	17			
Poole	92,111	4.5	Salisbury	35,800	30			

BRIGHTON 163,159

Hove	72,973	2.0	Cuckfield	22,070	17	-		
Peacehaven	4,786	5.0	Eastbourne	60,918	22			
Portstade	15,574	3.0	Crinstead	16,390	28			
Southwick	11,924	4.5	H'wards Heath	14,373	15			
			Horsham	23,250	23			
			Lewes	13,390	8			
			Littlehampton	17,060	20			
			Seaford	14,030	13			
			Worthing	80,329	11			

BRISTOL 437,048

Easton	4,130	5.0	Bath	80,901	13	-		
Filton	12,297	4.5	Burnham	10,480	27			
Hanham	4,963	4.0	Chippenham	18,510	22			
Keynsham	15,152	5.0	Weston	43,620	21			
Ashton	4,504	3.0						
Margetsfield	5,404	4.5						
Whilchurch	1,118	4.0						

BURNLEY 80,559

Accrington	39,018	5.0	Colne	19,030	6			
Brierfield	3,712	3.0	Rawtenshall	23,150	7			
Habergham	726	2.0	Todmorden	16,810	9	Blackburn	106,242	11
Hapton	1,713	3.0				Bradford	296,000	30
Nelson	2,519	4.0				Halifax	97,000	22
Padcham	9,899	3.0				Huddesfield	130,000	27
Scronstone	1,172	5.0				Leeds	510,000	40
						Manchester	661,000	24
						Preston	113,341	21

BURTON 50,751

Barton	2,468	5.0	Coalville	27,070	11	Birmingham	1,102,570	32
Cresley	5,258	4.0	Lichfield	18,130	13	Coventry	305,521	32
Hilton	968	5.0	Rugeley	15,140	18	Derby	132,408	11
Linton	1,958	4.5	Tamworth	16,120	17	Leicester	273,470	27
Newhall	7,492	3.0				Nottingham	311,899	28
Repton	1,850	4.5				Sheffield	494,344	35
Rolleston	2,162	3.0				Stoke	265,306	24
Swadlincote	6,471	4.0						
Tatenhall	681	3.0						
Tulbury	2,566	4.0						
Willington	1,318	5.0						

CAMBRIDGE 95,527

Barton	788	4.0	Bedford	63,334	29	-		
Colton	519	3.5	Bury	22,270	28			
Girton	3,115	3.0	Startford	20,490	26			
Granitchester	418	3.0	Ely	10,100	16			
Falbourn	2,906	4.0	Huntington	11,480	16			
Fen Dilton	674	3.5	Newmarket	11,460	13			
Histon	3,258	3.5						
Homingsea	355	3.5						
Landbeach	648	4.0						
Stow	447	4.0						
Tavesham	808	2.5						
Waten Beach	2,681	5.0						

CARDIFF								256,582
Lisvare	850	4.5	Abertillery	24,760	25	-		
Penarth	20,896	3.5	Bary	42,460	10			
St. Andrews	4,789	4.0	Bridgend	15,180	19			
			Aberdare	38,910	23			
			Llanlrissant	22,681	10			
			Maesteg	21,590	28			
			Merthyr	59,037	24			
			Newport	107,590	12			
			Portypridd	35,160	12			
			Portyport	39,000	22			
			Tredegar	19,620	25			

CARLISLE								67,798
Beaumont	404	4.0	Maryport	12,270	27	-		
Dalston	2,064	4.0	Penrith	10,870	18			
Orton	378	3.0						
Rockcliffe	747	5.0						
Wetherhall	3,852	4.5						

CHELTENHAM								72,154
Bishops Cleave	4,244	3.5	Cirencester	12,640	16	Birmingham	1,102,570	45
Churchdown	7,075	4.5	Evesham	12,980	16	Bristol	437,048	43
Gotherington	452	4.5	Gloucester	69,733	19	Oxford	106,291	42
Shurdington	1,648	3.0	Worcester	65,923	25	Swindon	91,159	32
Swindon	448	2.0						
Woodmancote	1,146	4.5						

CHESTER								59,268
Christteton	2,047	2.5	Crewe	53,195	24	Liverpool	747,000	18
Guildon	718	3.5	Ellesmere	48,200	7	Manchester	661,000	38
Littleton	522	2.5	Oswestry	11,940	28	Preston	113,341	48
Stanney	270	5.0	Northwick	19,460	18	Stoke	265,306	37
Mottington	481	2.5	Wrescham	36,300	12	Warrington	75,934	20
Trafford	459	3.5						
Eccleston	261	2.0						
Upton	7,708	2.5						
Saltney	3,709	2.0						

CHESTERFIELD								67,839
Brimington	8,163	2.5	Alfreton	22,830	11	Derby	132,408	26
Calow	2,768	2.0	Buxton	19,390	24	Doncaster	86,322	28
Clay Cross	9,163	5.0	Mansfield	53,218	12	Manchester	661,000	46
Dronfield	3,552	5.0	Matlock	19,390	10	Nottingham	311,899	26
North Wingfield	8,012	4.5	Relford	18,290	23	Sheffield	494,344	12
Unstone	2,576	4.0	Ripley	17,720	15	Stoke	265,306	45
			Worksop	35,400	15			

COLCHESTER

65,086

Abberton	177	3.5				Cambridge	95,527	48
Ardleigh	1,860	4.0				Ipswich	177,395	18
Alresford	926	5.0	Braintree	21,060	15			
Fingringhoe	590	3.5	Clacton	30,780	15			
Fordham	379	5.0	Harwich	14,150	19			
Lagenhoe	202	4.0	Wilham	10,190	14			
Lager	898	3.5						
Slanway	2,426	4.0						
West Hotte	1,294	3.0						
Wivenhoe	2,729	3.5						

COVENTRY

305,521

Ansty	234	5.0	Leamington	44,300	9	Birmingham	1,102,570	18
Bedworth	16,047	5.0	Harborough	12,300	28			
Corley	858	4.0	Nuneaton	57,376	8			
Kenilworth	14,449	5.0	Stratford	17,400	19			
Ryton	1,243	4.5	Rugby	51,968	12			
Shilton	695	5.0	Warwick	16,870	11			
Stoneleigh	5,409	4.0						
Wolston	1,436	5.0						

CREWE

53,195

Acton	293	5.0	Congleton	17,400	13	Chester	59,268	24
Barthomley	285	4.0				Liverpool	747,000	43
Minshill	340	4.0				Manchester	661,000	34
Hastington	2,879	2.0				Stoke	265,306	15
Nantwick	10,438	3.5				Warrington	75,934	26
Sandbach	9,862	5.0						
Warmingham	215	4.0						
Wistaton	4,519	1.5						
Willaston	1,818	3.0						

DARLINGTON

84,184

Aycliffe	12,098	5.0	Bishop Auckland	34,930	12	Middlesborough	157,000	15
Croft	1,992	3.0				Newcastle	265,000	33
Conniscliffe	321	2.5				Stockton	81,000	11
Middleton	2,337	4.0				York	104,392	48
Piercebridge	157	5.0						
Sandberge	557	3.5						
Walworth	182	4.5						

DERBY 132,408

Allestree	7,298	2.5	Alfreton	22,830	15	Birmingham	1,102,570	40
Alverton	13,855	3.0	Belper	15,760	8	Coventry	305,521	40
Breadsall	4,904	2.5	Burton	50,751	11	Leicester	273,470	28
Barrow	327	5.0	Fastwood	10,960	12	Nottingham	311,899	16
Chaddesden	15,622	2.5	Leek	19,100	28	Sheffield	494,344	37
Chellaston	3,702	4.0	Mansfield	53,218	24	Stoke	265,306	34
Duffield	3,375	5.0	Ripley	17,720	11			
Findem	1,059	4.5						
Mukleover	9,709	3.0						
Sponden	11,541	3.0						

DONCASTER 86,322

Adwick	3,359	4.0	Barnsley	74,704	12	Leeds	510,000	28
Bentley	6,208	2.0	Gainsborough	17,210	21	Nottingham	311,899	44
Carcroft	3,443	5.0	Goole	18,680	21	Sheffield	494,344	18
Carnborough	360	5.0	Hemsworth	14,550	12	York	104,392	34
Edenthorpe	2,824	3.5	Knottingley	13,320	17			
High Meltin	261	4.0	Pontefract	28,320	15			
Rossington	10,190	4.0	Scunthorpe	67,324	25			
Spotborough	7,469	2.5	Selby	10,670	20			
Wadworth	622	4.0	Thorne	15,280	10			
Warmsworth	2,959	3.0	Wakefield	61,628	18			

DUNDEE 182,978

-			Arbroath	20,063	17	-		
			Forfar	10,150	14			
			Montrose	10,783	30			
			Perth	41,497	22			

EASTBOURNE 60,918

East Dean	937	4.0	Lewes	13,390	17	Brighton	163,189	22
Jevington	188	3.5	Seaford	14,030	9	Hastings	66,478	17
Polegate	5,208	4.0						
Peversay	2,151	4.0						
Westham	1,932	4.0						
Willingdean	3,857	3.0						

EDINBURGH 468,361

-			Bathgate	13,467	19	Glasgow	979,978	44
			Falkirk	38,042	26			
			Grangemouth	20,425	25			
			Musselburgh	17,805	6			

EXETER 80,321

Clyst	2,043	5.0	Exmouth	20,810	11	Plymouth	204,409	42
Exminster	3,302	4.0	Newton	18,650	16			
Ide	598	2.0	Sidmouth	10,680	15			
Kenn	959	4.5	Taverton	13,456	15			
Pinhoe	3,451	3.0	Teignmouth	11,466	14			
Sowton	425	3.5	Torquay	54,066	23			
Stoke Cannon	336	4.0						
Shillingford	212	3.0						
Topsham	3,963	4.0						

GLOUCESTER 69,733

Brockworth	6,820	4.0						
Churchdown	7,075	3.5						
Hucclecote	4,598	2.5						
Longford	1,303	2.0						
Longlevens	6,966	3.0						
Maisemore	447	2.5	Cirencester	12,640	17	Bath	80,901	38
Minsterworth	497	4.5	Hereford	43,950	28	Bristol	437,048	35
Tibberton	274	5.0				Cheltenham	75,124	9
						Oxford	106,291	49
						Swindon	91,159	41

GREAT YARMOUTH 52,970

Caister	4,104	2.5	Lowestoft	42,450	10	Norwich	120,096	20
Filby	489	5.0						
Ormesby	321	5.0						
Scratby	1,357	5.0						

GREENOCK 74,560

Gourock	9,608	2.5	-			Glasgow	979,978	23
Inverkip	386	5.0						
Port Glasgow	22,559	3.0						

GRIMSBY 96,712

Barnetby	217	4.0	Louth	11,390	17	Lincoln	77,077	36
Coates	554	2.5	Scunthorpe	67,324	30			
Healing	1,296	3.5						
Holton	177	4.0						
Humberston	3,088	3.5						
Laceby	1,369	4.0						
Waltham	3,015	3.5						

HARROGATE								56,345
Hampwaithe	565	4.0	Ilkley	18,960	17			
Kittinghall	2,490	3.0	Otley	11,720	11			
Overblow	296	4.5	Ripon	10,760	11	Darlington	84,184	47
Knaresborough	9,309	3.5	Skipton	13,140	22	Doncaster	86,322	41
Ribston	183	5.0				Leeds	510,000	16
Rigton	365	5.0				York	104,392	22
Spotforth	286	4.5						

HASTINGS								66,478
Battle	4,517	5.0	Eastbourne	60,918	17	Brighton	163,159	37
Bexhill	289,411	5.0	Tunbridge Wells	41,280	28			
Fairlight	866	4.0						
Pelt	639	5.0						
Sedlescombe	991	5.0						
Westfield	1,779	4.5						

IPSWICH								117,395
Bramford	210	3.0	Bury	22,270	26			
Claydon	686	4.0	Colchester	65,086	18			
Kesgrove	3,382	4.0	Felixstowe	17,750	12			
Bealings	412	4.5						
Naiton	677	4.0						
Rushmere	2,024	2.5						
Tuddenham	356	3.0						
Washbrook	388	3.0						
Woolverstone	280	4.0						

KINGSTON-UPON-HULL								303,261
Bilton	2,380	5.0	Beverly	16,570	8	-		
Cottingham	11,294	4.0	Bridlington	26,250	29			
Paill	629	5.0	Gode	18,680	27			
Swine	222	5.0						

KIRKCALDY								52,390
Kinghorn	2,487	3.5	Cowdenbeath	11,438	9	-		
Burntisland	6,188	5.0	Dunfermline	49,555	13			
			Methil	20,600	8			
			Perth	41,497	28			
			St. Andrews	10,250	26			

LEICESTER 273,470

Ansley	3,784	4.5	Burton	50,751	26	Birmingham	1,102,570	39
Blaby	4,342	4.0	Grantham	25,670	30	Coventry	305,521	24
Birstall	10,143	3.0	Coalville	27,070	12	Nottingham	311,899	25
Enderly	4,024	4.0	Loughborough	39,270	12			
Kirkby	4,931	4.0	Harborough	12,300	15			
Narborough	3,479	5.0	Melton	10,850	15			
Oadby	12,256	3.5	Nuneaton	57,376	18			
Rotheby	3,033	5.0						
Syston	6,455	5.0						
Thurlnaston	7,160	3.5						
Thurnby	2,722	4.0						
Thurslaston	378	5.0						
Whetston	1,466	4.5						
Wigston	36	3.5						

LINCOLN 77,077

Bracebridge	2,825	2.5	Gainsborough	17,210	18	Grimsby	96,712	36
Branston	2,035	4.0	Grantham	25,670	25	Sheffield	494,344	46
Burton	165	2.0	Newark	24,780	16			
Heighington	763	4.0	Retford	18,290	28			
Nettleham	1,940	4.0						
Hykeham	5,308	4.0						
Reepham	704	4.5						
Saxby	66	5.0						
Skellingthorpe	1,328	3.5						
Waddington	4,193	4.5						
Washingborough	1,141	3.0						

MANSFIELD 53,218

Blidworth	7,308	4.5	Alfreton	22,830	10	Chesterfield	67,839	12
Hathwaithe	4,521	4.5	Matlock	19,390	17	Derby	132,408	24
Measley	2,754	3.5	Newark	24,780	18	Doncaster	86,322	30
Shirebrook	11,635	4.0	Redford	18,290	23	Nottingham	311,899	14
Sutton	13,940	3.0	Worksop	35,400	13	Sheffield	494,344	23

MERTHYR TYDFILL 59,032

-			Aberdare	38,910	7	Cardiff	256,586	24
			Abertillery	24,760	15	Newport	108,123	28
			Pontypridd	35,160	12	Swansea	167,322	30
			Pontyport	39,000	22			
			Tredegar	19,620	7			

NEWPORT 108,123

Caerlean	1,373	2.5	Aberdare	38,910	27	Cardiff	256,582	12
Rogerstone	5,931	4.0	Abertillery	24,760	16			
			Merthyr	58,310	28			
			Pontypridd	35,160	18			
			Pontypool	39,000	10			
			Tredegar	19,620	22			

NORTHAMPTON 105,421

Blisworth	1,192	5.0	Bedford	63,334	21	Birmingham	1,102,705	50
Brafield	611	4.5	Bletchley	20,610	23	Coventry	305,521	32
Broughton	2,300	3.5	Barnsbury	23,080	25	Leicester	273,470	32
Brampston	230	4.0	Kettering	38,840	14	Luton	131,000	35
Duston	4,885	2.5	Leamington	44,300	25			
Ecton	461	5.0	Leighton	12,880	27			
Harpole	1,097	4.0	Harborough	12,300	18			
Kislingbury	806	3.5	Rugby	51,968	20			
Milton	658	3.5	Rushden	17,490	15			
Moulton	3,424	4.5	Warwick	16,870	27			
Pilsford	822	4.5	Wellingborough	31,910	10			
Rothesthorpe	230	3.5						
West Fawell	5,105	2.0						

NORWICH 120,096

Calton	2,592	2.5	Great Yarmouth	52,970	20	Ipswich	177,395	43
Cringleford	1,124	3.0	Lowestoft	42,450	27			
Drayton	1,346	4.5						
Farningham	436	4.5						
Hellesden	9,744	2.5						
Honham	1,361	4.5						
Spoxworth	1,128	4.0						
Sprouston	9,609	2.5						
Swardeston	447	4.5						
Thorpe	10,788	2.5						

NOTTINGHAM 311,899

Beeston	19,487	4.0	Burton	50,751	27			
Burton	2,447	5.0	Derby	132,408	16			
Lambley	979	4.5	Eastwood	10,960	8			
Radcliffe	82	4.5	Grantham	25,670	24			
Ruddington	5,158	5.0	Ilkeston	34,990	7	Birmingham	1,102,570	50
Stappleford	13,307	4.0	Loughborough	39,270	15	Sheffield	494,344	37
			Mansfield	53,218	14			
			Melton	16,850	18			
			Newark	24,780	20			

NUNEATON 57,376

Arley	3,319	5.0	Tamworth	16,120	18	Birmingham	1,102,570	22
Atherstone	5,453	5.0				Coventry	305,521	8
Bedworth	16,847	3.0				Leicester	27,470	18
Higham	773	3.0						
Hartshill	2,583	3.0						
Hinckley	41,608	4.5						

OXFORD

106,291

Cassington	532	3.0	Abingdon	15,680	7	Reading	125,350	27
Cumnor	4,197	3.5	Aylesbury	32,510	23			
Eynsham	2,357	5.0	Banbury	23,080	23			
Islip	620	4.0	Swindon	91,159	29			
Horspath	1,540	3.5						
Kennington	3,452	2.5						
Kidlington	8,514	4.0						
Littlemore	8,529	3.0						
Hinskey	4,452	1.5						
Sandford	813	3.0						
Stanton	438	4.5						
Wheatley	2,208	5.0						
Yarnton	1,371	3.0						

PETERBOROUGH

63,340

Ailsworth	942	4.5	Marsh	13,240	18	Cambridge	95,527	35
Eye	2,364	3.0	Spalding	15,180	19	Leicester	273,470	41
Farcet	1,236	2.5	Stanford	12,560	14	Northampton	105,421	40
Glinton	1,067	4.5	Wisbeck	17,520	21			
Markolm	205	3.0						
Newborough	719	4.0						
Old Fletton	190	5.0						

PLYMOUTH

204,409

Bickleigh	957	5.0	-	-	
Plympton	12,396	5.0			
Plymstock	14,700	4.0			
Saltash	7,425	2.0			
Torpoint	4,268	2.5			

PORTSMOUTH

215,077

Fareham	24,820	5.0	Chichester	20,280	17	-
Gosport	62,457	1.0	Fareham	68,690	8	
Havant	11,596	5.0	Haslemere	13,210	29	
Portchester	12,178	4.0				
Purbrock	6,717	4.0	Bognor Regis	29,260	24	
			Littlehampton	17,060	30	
			Southampton	204,827	21	
			Winchester	30,310	27	

PRESTON

113,341

Barton	1,497	5.0	Blackburn	106,241	10	Blackpool	153,185	17
Farington	4,434	2.5	Chorley	31,060	9	Liverpool	747,000	30
Fulwood	7,229	2.0	Clitheroe	12,550	17	Manchester	661,000	30
Grimsargh	835	4.5	Southport	82,000	18			
Hutton	1,720	3.0						
Lea Town	3,657	4.0						
Langton	3,884	4.5						
Leyland	19,413	5.0						
Walton	5,104	2.5						
Woodplumpton	1,694	4.0						

RUGBY 51,968

Barby	427	4.5	Harborough	12,300	18	Birmingham	1,102,570	30
Brincklow	1,092	5.0				Coventry	305,521	12
Culthorpe	122	3.5				Leicester	273,470	21
Lawford	449	3.5				Northampton	105,421	20
Churchover	647	3.5				Oxford	106,921	48
Kilsby	227	4.5						
Milton	2,761	2.5						
Poulton	440	2.5						
Willoughby	302	5.0						

SCUNTHORPE 67,324

Burningham	986	3.5	Goole	18,680	21	Doncaster	86,322	23
Burton	1,398	5.0				Grimsby	96,712	30
Flexborough	449	3.0				Lincoln	77,072	28
Keadby	2,007	5.0						
Roxby	416	4.0						

SHEFFIELD 494,344

Barnsworth	5,344	4.0	Barnsley	74,704	14	Leeds	510,000	33
Catcliffe	2,030	4.0	Buxton	219,390	28	Manchester	661,000	38
Kimberworth	9,650	5.0	Chesterfield	67,839	12			
Orgreave	504	4.5	Doncaster	86,322	18			
Treeton	2,289	4.0	Holmforth	18,620	21			
			Huddersfield	132,270	26			
			Maltby	14,090	13			
			Maltock	19,930	22			
			Mexborough	16,690	13			
			Retford	18,290	23			
			Rotherham	86,510	6			
			Worksop	35,400	17			

SOUTHAMPTON 204,822

Burlesdon	3,560	4.0	Basingstoke	30,360	29	Bournemouth	154,296	30
Chandlers	9,058	5.0	Lymington	30,310	18	Portsmouth	215,077	21
Eastleigh	27,584	5.0	Salisbury	35,800	23			
Hamble	3,007	4.5	Winchester	30,310	11			
Hedge End	4,464	4.0						
Netley	2,618	3.0						
West End	5,064	4.0						

SOUTHPORT 82,000

Halsall	1,944	5.0	Chorley	31,060	20	Liverpool	747,000	20
Scansbath	2,895	3.5	Ormskirk	24,350	8	Manchester	661,000	39
						Preston	113,341	17

STOKE

265,306

Barleston	2,459	4.5	Buxton	19,390	29	Birmingham	1,102,570	45
Cayerswall	10,313	5.0	Congleton	17,400	13	Manchester	661,000	38
Newcastle	25,688	2.0	Crewe	53,195	15	Sheffield	434,344	47
Whitmere	645	5.0	Leek	19,100	11			
			Macclesfield	40,210	20			
			Nantwich	10,980	17			
			Sandbach	10,350	14			
			Stafford	49,480	16			

SUNDERLAND

189,686

Boldon	10,134	4.5	Auckland	34,960	23	Middlesborough	157,000	27
Cleadon	3,782	3.0	Chester	19,380	10	Newcastle	265,000	12
Herrington	2,153	5.0	Durham	23,050	13	Stockton	81,000	26
Houghton	9,586	5.0	South Shields	108,770	7			
Ryhope	9,786	3.5	West H'pool	77,035	19			
Washington	2,520	5.0						
Whitburn	4,763	3.0						

SWANSEA

167,322

Gorsienan	3,701	5.0	Aberdare	38,910	27	Cardiff	256,582	40
Llanfellach	2,650	3.5	Bridge Ford	15,180	21			
			Carmarthen	12,820	27			
			Llanelly	29,270	11			
			Maesteg	21,590	30			
			Neath	30,250	8			
			Port Talbut	51,750	9			

SWINDON

91,159

Blunsden	1,444	3.5	Chipperham	18,510	20	Bristol	437,048	41
Chistleton	2,598	4.0	Newsbury	21,380	25	Oxford	106,291	29
Haydon	564	2.5				Reading	125,390	38
Liddington	364	4.0						
Lydiard	958	3.5						
Purton	3,295	4.0						
Stratton	11,191	2.5						
Warborough	972	4.0						
Wootton	4,390	5.0						
Wroughton	5,108	3.0						

TORQUAY

54,066

Marledon	1,179	3.5	Brixham	11,390	8	Exeter	80,321	22
Newton	18,060	5.0	Teignmouth	11,640	9	Plymouth	204,409	32
Paignton	30,292	2.5						

WAKEFIELD 61,628

Crigglestone	5,147	3.0	Batley	40,270	7	Barnsley	74,704	10
Crofton	5,140	3.5	Knottingdean	13,320	13	Halifax	97,000	17
Harbury	8,642	3.0	Hemsworth	14,550	8	Huddersfield	130,000	13
Normarton	6,024	3.5	Pontefract	28,320	9	Leeds	510,000	9
Osset	14,737	3.0				Sheffield	434,344	24
Outewood	4,848	2.5				York	104,392	29
Sharleston	2,967	3.5						
Stanley	3,059	3.0						
Walton	1,724	2.5						
Wrenthorpe	5,356	2.0						

WARRINGTON 75,954

Appleton	4,656	4.0	Altringham	41,250	12	Blackburn	106,242	30
Burton	2,766	4.0	Chester	59,268	20	Liverpool	747,000	17
Grapperhall	7,746	2.5	Congleton	17,400	27	Manchester	661,000	18
Halton	1,467	5.0	Crewe	53,195	28	Preston	113,344	29
Lymm	1,955	5.0	Knutsford	10,100	12	Stoke	265,306	37
Stratham	1,456	4.5	Northwich	19,460	12			
Stretton	579	3.5	Sandbach	10,350	25			
Stockton	6,684	2.0	Widnes	52,186	7			
Woolston	3,434	4.0						

WEST HARTLEPOOL 77,035

Elwick	260	3.5	Billinghorn	33,490	9	Middlesborough	157,000	9
Greathorn	1,416	3.0	Bishop A'Land	34,960	21	Newcastle	265,100	28
Hartlepool	17,675	2.0	Durham	23,050	18	Stockton	81,000	11
						Sunderland	189,000	19

WIDNES 52,186

Cronton	737	2.5	Northwich	19,460	12	Chester	59,268	17
Frodesham	5,661	4.0				Liverpool	747,000	13
Sankey	5,827	4.5				Warrington	75,954	7
Prescott	13,079	5.0						
Preston	417	4.5						
Penketh	5,220	3.5						
Rainhill	7,913	4.0						
Runcorn	26,038	1.0						

WIGAN 78,690

Abram	2,353	2.5	Atherton	19,520	7	Blackburn	106,242	19
Adlington	1,303	4.5	Chorley	31,060	8	Bolton	161,000	10
Ashton	2,744	5.0	Leigh	46,500	7	Liverpool	747,000	20
Billinge	6,079	5.0	Ormskirk	24,520	12	Manchester	661,000	19
Boars	2,568	2.5	Warrington	78,690	12	Preston	113,342	17
Orrd	4,951	3.5				Southport	82,000	20
Shevington	4,887	3.5						
Standish	1,020	2.0						
Up Holland	4,664	3.5						
West Houghton	2,021	4.5						

WORCESTER 65,923

Brodicot	28	2.5	Bromsgrove	36,970	13	Birmingham	1,102,570	27
Broughton	136	5.0	Evesham	12,980	16	Cheltenham	75,124	25
Crowle	480	4.5	Great Malvern	28,030	8	Gloucester	69,723	27
Hallow	1,204	2.5	Hereford	43,950	26			
Kempsey	1,704	3.5						
Leigh	979	5.0						
Tibberton	321	3.5						

WORTHING 80,329

Angmering	2,856	5.0	Bognor Regis	29,626	16	Brighton	163,159	11
Clapham	320	4.0	Chichester	20,280	23	Portsmouth	215,077	38
East Preston	4,039	5.0	Horsham	23,250	20			
Ferring	3,449	3.5	Littlehampton	17,060	9			
Findon	1,289	4.5						
Kingston	378	4.0						
Lancing	13,319	3.0						
Patching	272	4.0						
Shoreham	17,410	5.0						
Sompting	5,830	2.5						
Steyning	2,695	5.0						

YORK 104,392

Bryam	432	4.0	Beverley	16,350	29	Kingston	303,261	38
Bishopthorpe	1,263	3.0	Goole	18,680	25	Leeds	510,000	24
Conmanthorpe	1,027	4.0	Harrogate	56,345	22	Stockton	81,000	48
Dunnington	983	4.0	Ripon	10,760	24			
Haxby	2,407	4.0	Selby	10,670	14			
Heslington	1,233	2.0	Thorne	18,280	29			
Hotby	111	4.5						
Hunlington	6,435	3.0						
Nieburn	433	4.0						
Poppleton	597	3.0						
Rearswick	2,001	2.0						
Rufforth	265	2.5						
Skelton	944	3.5						
Stockton	859	4.0						

TABLE A2

Components of the Tertiary Sector - 1961

KEY: Foot of Table

UNIT ;No.Employed per 10;000

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Aberdeen	152	1,004	2,166	237	1,403	1,087	535	6,584
Barnsley	151	720	1,834	136	934	931	479	5,185
Barrow	210	719	1,190	116	677	748	439	4,099
Bath	456	541	1,869	180	1,402	1,187	1,458	7,093
Bedford	317	664	1,652	236	1,341	979	538	5,727
Blackburn	219	575	1,222	131	735	457	307	3,846
Blackpool	198	736	2,099	209	816	2,692	588	7,338
Bournemouth	183	771	2,552	589	1,015	2,351	488	7,949
Brighton	225	1,004	1,871	510	1,164	1,564	572	6,910
Bristol	178	1,185	1,810	308	1,081	1,036	421	6,019
Burnley	227	608	1,351	178	755	799	316	4,234
Burton	203	566	1,372	206	810	615	401	4,173
Cambridge	189	798	1,634	324	2,276	1,148	640	7,009
Cardiff	182	1,152	1,752	302	919	1,187	713	6,207
Carlisle	260	1,490	1,648	218	1,009	1,221	633	6,479
Cheltenham	207	593	1,830	328	1,228	1,312	1,346	6,844
Chester	248	1,119	1,942	350	1,323	1,370	1,089	7,441
Chesterfield	148	569	1,546	182	1,003	901	364	4,713
Colchester	166	689	1,435	154	1,371	1,020	1,644	6,479
Coventry	121	316	969	105	555	582	278	2,926
Crewe	121	2,371	1,035	111	514	549	356	5,057
Darlington	227	1,135	1,448	148	741	870	363	4,932
Derby	207	805	1,133	122	648	609	395	3,919
Doncaster	170	1,118	1,467	184	814	906	429	5,088
Dundee	147	699	1,640	211	953	832	381	4,863
Eastbourne	341	585	1,957	272	1,674	2,618	476	7,923
Edinburgh	160	949	1,688	417	1,351	1,221	751	6,517
Exeter	219	1,225	1,909	491	1,731	1,166	870	7,611
Gloucester	290	1,185	1,866	216	1,040	863	1,032	6,492
Great Yarmouth	289	654	1,879	185	943	1,228	369	5,547
Greenock	145	947	1,385	121	906	661	706	4,871
Grimsby	199	1,221	2,104	163	805	915	463	5,870
Harrogate	244	851	1,793	713	1,182	1,954	855	7,592
Hastings	214	774	2,091	236	1,671	1,986	678	7,650
Ipswich	248	773	1,763	265	913	976	454	5,372

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Kingston	168	1,544	1,877	190	779	806	378	5,742
Kirkcaldy	99	695	1,143	126	1,065	1,160	501	4,789
Leicester	143	457	1,424	207	794	659	245	3,929
Lincoln	163	862	1,440	226	912	849	679	5,131
Mansfield	273	613	1,718	107	961	939	436	5,047
Merthyr	165	687	1,295	85	790	598	531	4,151
Newport	220	1,424	1,410	175	810	909	502	5,450
Northampton	223	712	1,668	299	850	783	444	4,979
Norwich	234	759	1,721	452	842	1,006	473	5,487
Nottingham	193	638	1,678	272	868	888	415	4,952
Nuneaton	100	758	1,434	87	956	773	430	4,538
Oxford	92	566	1,269	158	1,906	950	384	5,325
Peterborough	146	1,157	1,575	202	577	911	359	4,927
Plymouth	243	763	1,636	214	772	875	2,508	7,011
Portsmouth	313	483	1,527	192	949	915	1,964	6,343
Preston	215	1,025	1,663	202	756	734	678	5,273
Rugby	133	550	1,107	111	867	624	255	3,647
Scunthorpe	70	437	1,069	115	580	502	236	3,009
Sheffield	185	569	1,373	146	820	699	280	4,072
Southampton	256	1,851	1,833	269	828	1,061	510	6,362
Southport	180	651	2,096	252	1,073	1,835	677	6,764
Stoke	146	587	1,164	134	639	575	245	3,490
Sunderland	246	539	1,786	153	825	852	350	4,751
Swansea	299	1,357	1,855	230	1,088	954	617	6,400
Swindon	172	1,453	1,583	132	670	717	377	5,099
Torquay	275	596	2,486	312	1,083	2,711	670	8,133
Wakefield	311	742	1,374	147	1,311	832	1,006	5,723
Warrington	217	697	1,214	127	566	624	311	3,756
West H'pool	159	974	1,598	162	909	858	404	5,064
Widnes	71	568	842	90	554	568	374	3,067
Wigan	235	780	2,002	183	788	728	345	5,061
Worcester	272	758	2,001	167	989	940	678	5,805
Worthing	126	597	2,447	402	1,272	1,949	843	7,636
York	171	1,449	1,522	266	951	921	625	5,905

KEY - Industrial order in brackets

(1) = Gas, electricity and water (18)
(2) = Transport and communication (19)
(3) = Distributive Trades (20)
(4) = Insurance, banking and finance (21)
(5) = Professional and scientific (22)
(6) = Miscellaneous services (23)
(7) = Public administration (24)
(8) = Total of industrial orders 18 to 24

TABLE A3

Size and Changes in Tertiary Sector, 1931 and 1951,

Retail Output 1961

KEY: Foot of Table

UNITS; see key	(1)	(2)	(3)	(4)	(5)
Aberdeen	5,250	6,416		168	0.213
Barnsley	3,454	4,670	50.1	515	0.254
Barrow	2,947	3,853	39.0	246	0.158
Bath	5,667	6,882	25.1	211	0.238
Bedford		5,660		67	0.261
Blackburn	4,485	3,679	-14.3	167	0.179
Blackpool	6,667	7,291	10.0	47	0.238
Bournemouth	6,572	7,984	20.9	-35	0.277
Brighton	6,522	6,990	5.9	-80	0.240
Bristol	5,316	5,634	13.2	385	0.187
Burnley	3,704	3,670	14.3	564	0.172
Burton	3,677	4,062	13.5	111	0.221
Cambridge		7,124		-115	0.267
Cardiff	6,654	6,144	-6.7	63	0.201
Carlisle	5,782	5,513	12.0	966	0.257
Cheltenham		6,581		263	0.238
Chester	5,885	7,179	26.4	262	0.333
Chesterfield	3,338	4,254	41.1	459	0.234
Colchester		6,091		388	0.225
Coventry	2,464	2,694	18.7	232	0.175
Crewe		4,440		617	0.197
Darlington	4,172	4,855	18.1	77	0.216
Derby	4,397	3,870	-10.9	49	0.268
Doncaster	4,426	3,401	15.0	1,687	0.299
Dundee	3,919	4,243		640	0.181
Eastbourne	6,856	8,006	15.5	-83	0.274
Edinburgh	5,250	6,327	10.1	190	0.213
Exeter	6,209	6,970	22.5	641	0.248
Gloucester	5,723	6,765	13.4	-273	0.259
Great Yarmouth	6,341	6,064	-12.5	-517	0.248
Greenock	3,100	4,291	15.3	580	0.145
Grimsby	4,834	5,895	21.4	-25	0.202
Harrogate		8,157		-565	0.257
Hastings	6,909	8,038	10.7	-388	0.199
Ipswich	4,967	5,414	8.1	-42	0.221

	(1)	(2)	(3)	(4)	(5)
Kingston	5,757	5,799	-0.3	-57	0.178
Kirkcaldy	3,267	4,449	11.5	340	0.204
Leicester	3,858	3,775	1.8	154	0.235
Lincoln	4,561	4,929	12.4	202	0.234
Mansfield		4,844		203	0.264
Merthyr	3,233	4,363	28.3	-212	0.132
Newport	5,792	5,638	-5.9	-188	0.207
Northampton	4,070	4,403	22.3	576	0.231
Norwich	4,782	5,205	14.7	282	0.262
Nottingham	4,459	4,567	11.0	383	0.223
Nuneaton		4,193		345	0.199
Oxford	5,123	5,640	3.9	-315	0.280
Peterborough		5,161		-234	0.236
Plymouth	6,580	6,371	6.5	640	0.191
Portsmouth	6,619	6,641	-4.2	-298	0.216
Preston	4,971	4,335	6.1	538	0.232
Rugby		4,270		-623	0.182
Scunthorpe		2,942		67	0.177
Sheffield	3,632	3,885	12.1	187	0.180
Southampton	6,218	6,468	1.3	-166	0.223
Southport	6,388	6,986	5.9	-222	0.213
Stoke	2,899	3,043	20.4	447	0.166
Sunderland	4,503	4,269		482	0.179
Swansea	5,379	5,539	19.0	861	0.182
Swindon	3,077	4,729	65.7	370	0.256
Torquay		7,932		151	0.241
Wakefield	4,610	5,700	24.1	23	0.235
Warrington	3,129	3,678	20.0	78	0.223
West H'pool	5,437	4,936	-6.9	128	0.172
Widnes		2,810		257	0.134
Wigan	3,801	4,677	33.1	384	0.237
Worcester	5,207	5,577	11.4	228	0.245
Worthing		7,676		-40	0.246
York	5,810	5,808	1.6	100	0.235

KEY

(1) = ~~percentage~~ employment in tertiary sector 1931 (per 100000)
(2) = ~~percentage~~ employment in tertiary sector 1951 (;;)
(3) = change since 1931 as percentage of size of tertiary sector 1961
(4) = change in percentage employment in tertiary sector 1951 - 1961
(5) = per capita retail output (£00's)

TABLE A4

Rateable Values and Income Proxy Variables

KEY: Foot of Table

UNITS; see key

	(1)	(2)	(3)	(4)	(5)
Aberdeen	9.3	4.02	0.422	0.220	6373
Barnsley	16.6	6.93	0.745	0.239	6255
Barrow	18.1	4.95	0.485	0.293	7495
Bath	15.9	4.37	0.629	0.209	7859
Bedford	31.2	6.95	0.735	0.261	4304
Blackburn	15.9	5.50	0.675	0.275	7902
Blackpool	24.5	8.13	0.581	0.195	5557
Bournemouth	26.0	3.67	0.791	0.162	4773
Brighton	24.9	3.97	0.573	0.212	6793
Bristol	23.9	4.95	0.464	0.237	6265
Burnley	17.8	5.73	0.704	0.279	6668
Burton	28.0	6.29	0.725	0.257	5936
Cambridge	29.2	5.13	0.701	0.216	6853
Cardiff	25.3	6.68	0.456	0.232	5465
Carlisle	18.9	4.98	0.557	0.218	7711
Cheltenham	22.5	4.57	0.501	0.235	5722
Chester	25.2	8.93	0.831	0.232	0
Chesterfield	23.9	5.33	0.733	0.232	6127
Colchester	18.6	4.79	0.607	0.266	6586
Coventry	21.0	7.53	0.350	0.256	8748
Crewe	15.5	7.47	0.646	0.313	6987
Darlington	23.8	5.78	0.557	0.272	5730
Derby	33.3	6.28	0.924	0.308	6842
Doncaster	24.1	5.73	0.751	0.280	6135
Dundee	10.1	4.27	0.345	0.256	7551
Eastbourne	19.4	3.81	0.589	0.148	4631
Edinburgh	12.4	4.53	0.345	0.226	5647
Exeter	31.9	3.65	0.682	0.180	3913
Gloucester	21.4	6.62	0.906	0.244	6246
Great Yarmouth	26.1	4.33	0.446	0.228	4752
Greenock	8.1	5.31	0.372	0.267	7488
Grimsby	18.5	6.70	0.508	0.234	7206
Harrogate	15.3	4.19	0.475	0.184	6478
Hastings	13.5	3.11	0.488	0.142	6625
Ipswich	22.3	4.44	0.752	0.249	6713

	(1)	(2)	(3)	(4)	(5)
Kingston	16.5	5.65	0.405	0.247	4974
Kirkcaldy	11.6	4.87	0.549	0.192	6829
Leicester	28.1	4.49	0.590	0.266	8222
Lincoln	16.2	5.18	0.656	0.275	7194
Mansfield	19.7	7.12	0.724	0.226	4772
Merthyr	12.5	5.05	0.269	0.271	8164
Newport	24.1	9.58	0.122	0.250	8426
Northampton	21.2	4.31	0.767	0.364	8442
Norwich	25.1	3.75	0.748	0.242	7203
Nottingham	27.8	4.40	0.694	0.260	6123
Nuneaton	17.4	5.17	0.617	0.268	6900
Oxford	34.5	6.08	0.689	0.305	7653
Peterborough	17.5	5.13	0.752	0.288	7179
Plymouth	19.0	3.84	0.480	0.226	5138
Portsmouth	18.1	4.83	0.512	0.235	6018
Preston	22.1	6.47	0.673	0.270	5435
Rugby	19.4	8.28	0.560	0.285	8271
Scunthorpe	56.4	9.98	0.513	0.275	8446
Sheffield	25.8	9.64	0.411	0.288	7162
Southampton	27.1	7.56	0.567	0.258	4004
Southport	17.1	4.17	0.471	0.232	2252
Stoke	21.6	7.11	0.554	0.247	7470
Sunderland	17.6	4.49	0.432	0.430	6933
Swansea	20.3	7.23	0.435	0.237	6627
Swindon	18.5	5.36	0.525	0.271	6872
Torquay	21.0	3.71	0.671	0.130	5066
Wakefield	20.7	4.59	0.796	0.249	6209
Warrington	27.2	8.88	0.723	0.281	7535
West H'pool	14.0	4.99	0.671	0.252	7230
Widnes	20.8	8.65	0.589	0.269	8510
Wigan	19.7	4.28	0.701	0.245	5703
Worcester	25.1	4.82	0.719	0.256	5485
Worthing	17.1	3.50	0.549	0.108	0
York	17.7	4.67	0.555	0.244	9052

KEY

(1) = Non-domestic rateable value (per capita). (£ per capita)
(2) = Electricity consumption (per capita) (kw ;; ;;)
(3) = Television licences in force (per household)
(4) = Index of earnings of secondary sector (see text)
(5) = Index of 'exporting' sector (see text)

TABLE A5

Population Change to 1961

KEY: Foot of Table

UNITS; see key	(1)	(2)	(3)	(4)	(5)
Aberdeen	17.3	11.7	13.1	8.4	1.5
Barnsley	45.1	32.3	3.3	-1.6	-1.2
Barrow	11.3	1.8	-14.3	-1.9	-3.9
Bath	18.5	14.5	15.2	15.0	2.0
Bedford	38.2	36.5	33.9	32.7	16.2
Blackburn	-21.6	-25.2	-19.5	-15.6	-4.7
Blackpool	-9.8	-18.0	30.5	27.8	3.9
Bournemouth	61.2	42.7	37.9	24.3	6.1
Brighton	24.4	19.6	9.7	12.8	4.1
Bristol	22.5	18.3	11.1	6.0	-1.3
Burnley	-20.8	-32.5	-28.0	-22.0	-5.5
Burton	0.7	4.9	3.7	2.4	3.2
Cambridge	47.3	41.6	35.7	26.6	14.7
Cardiff	29.0	22.0	12.7	10.4	5.1
Carlisle	25.5	23.0	22.3	15.5	0.1
Cheltenham	32.5	32.2	32.1	31.5	12.9
Chester	34.9	34.2	6.1	1.5	4.5
Chesterfield	29.3	18.5	9.7	5.4	-1.0
Colchester	41.1	33.3	33.1	24.5	11.7
Coventry	70.1	65.2	46.4	38.1	15.5
Crewe	21.0	16.0	6.5	5.4	1.5
Darlington	46.0	31.9	20.6	14.4	-1.0
Derby	13.3	6.8	0.7	-7.6	-6.7
Doncaster	54.3	43.9	35.8	23.9	5.0
Dundee	12.0	8.9	7.2	3.2	3.1
Eastbourne	28.5	13.8	-5.0	2.1	5.1
Edinburgh	18.0	13.5	10.0	6.3	0.4
Exeter	31.0	29.0	21.1	13.4	6.0
Gloucester	31.2	28.3	18.7	15.6	3.5
Great Yarmouth	3.2	-5.5	-14.6	-7.2	3.5
Greenock	8.6	-0.7	-8.9	-6.0	-2.3
Grimsby	-18.4	-40.6	21.3	4.3	4.3
Harrogate	49.6	40.2	17.7	15.3	10.5
Hastings	1.5	8.1	-1.5	0.4	1.4
Ipswich	43.3	37.1	30.5	23.5	8.4

	(1)	(2)	(3)	(4)	(5)
Kingston	20.8	8.4	4.1	-0.1	1.4
Kirkcaldy	35.0	24.5	14.7	15.5	6.4
Leicester	17.0	14.5	6.5	-0.1	-4.3
Lincoln	33.3	20.5	14.3	14.1	10.0
Mansfield	59.8	30.7	16.5	13.4	3.5
Merthyr	-37.1	-35.7	-35.7	-20.4	-3.5
Newport	37.8	22.6	4.4	0.4	2.4
Northampton	17.5	14.4	10.6	6.3	0.1
Norwich	5.2	-1.1	0.5	-5.1	-1.2
Nottingham	23.2	16.7	12.6	8.6	1.3
Nuneaton	72.6	56.4	26.6	18.9	5.2
Oxford	53.6	50.1	36.7	24.2	7.2
Peterborough	51.3	47.0	29.1	18.2	-1.0
Plymouth	5.5	-1.4	-7.0	-6.6	-1.7
Portsmouth	11.5	-8.5	-16.7	-18.7	-8.6
Preston	0.4	-3.3	-5.9	-7.5	-7.2
Rugby	67.7	58.2	14.3	12.6	12.6
Scunthorpe	83.8	71.3	51.4	41.9	21.4
Sheffield	16.9	18.8	-5.6	-6.1	-3.7
Southampton	37.6	29.2	21.4	14.1	12.9
Southport	22.5	15.1	6.6	3.7	-0.1
Stoke	19.1	11.6	-0.1	-4.3	-3.7
Sunderland	23.0	21.4	3.9	2.0	4.3
Swansea	14.0	5.9	5.8	1.5	3.8
Swindon	50.7	44.4	37.6	31.5	24.4
Torquay	60.9	40.1	26.4	14.3	1.5
Wakefield	21.7	16.5	13.7	4.8	2.1
Warrington	15.5	5.0	-6.2	-10.0	-6.2
West H'pool	12.8	17.1	8.4	8.6	5.7
Widnes	38.4	32.4	25.5	22.2	6.6
Wigan	-4.7	-13.3	-13.6	-8.4	-7.5
Worcester	25.5	26.0	23.1	20.1	9.4
Worthing	72.0	62.3	53.8	42.1	25.7
York	25.3	20.8	6.1	1.3	-1.0

KEY

(1) = Population change 1901 - 1961 as percentage of 1961 figure
(2) = Population change 1911 - 1961 as percentage of 1961 figure
(3) = Change in population 1921 - 1961 as percentage of 1961 figure
(4) = Change in population 1931 - 1961 as percentage of 1961 figure
(5) = Change in population 1951 - 1961 as percentage of 1961 figure

TABLE A6

Miscellaneous Independent Variables - 1961

KEY: Foot of Table

UNITS, see key	(1)	(2)	(3)	(4)	(5)
Aberdeen	45.47	0.210	3.411	4.165	3
Barnsley	45.12	0.574	3.522	4.570	-1
Barrow	44.83	0.739	3.141	4.216	2
Bath	45.72	0.006	3.287	4.533	6
Bedford	46.53	0.274	8.158	11.844	2
Blackburn	53.12	0.246	3.037	4.307	-2
Blackpool	45.65	0.045	3.485	4.347	5
Bournemouth	40.94	0.067	3.300	4.262	33
Brighton	45.51	0.051	3.667	4.986	6
Bristol	45.98	0.093	3.414	4.722	2
Burnley	52.84	0.405	2.999	4.233	0
Burton	44.90	0.095	3.417	4.797	1
Cambridge	43.50	0.016	3.847	5.169	2
Cardiff	44.23	0.336	3.439	4.599	1
Carlisle	48.60	0.291	3.343	4.564	2
Cheltenham	46.84	0.030	3.462	4.913	11
Chester	46.90	0.240	3.304	4.214	4
Chesterfield	45.82	0.223	3.317	4.334	0
Colchester	45.23	0.180	3.712	4.803	2
Coventry	49.98	0.112	3.771	5.409	0
Crewe	46.71	0.004	3.354	4.661	0
Darlington	46.33	0.356	3.290	4.339	0
Derby	47.54	0.228	3.660	5.164	2
Doncaster	45.03	0.286	3.667	4.736	3
Dundee	51.15	0.402	3.059	4.199	1
Eastbourne	42.43	0.013	7.006	9.424	12
Edinburgh	51.74	0.145	3.341	4.588	21
Exeter	44.55	0.067	3.377	4.268	7
Gloucester	46.27	0.103	3.424	4.548	0
Great Yarmouth	43.76	0.107	3.414	4.391	2
Greenock	42.91	0.519	3.319	4.192	-1
Grimsby	44.57	0.175	3.278	4.510	2
Harrogate	45.15	0.225	2.903	4.486	4
Hastings	39.48	0.077	3.101	4.137	2
Ipswich	43.41	0.133	3.565	4.588	1

	(1)	(2)	(3)	(4)	(5)
Kingston	44.48	0.156	3.351	4.523	-2
Kirkcaldy	43.78	0.196	3.510	4.307	0
Leicester	52.53	0.330	3.573	5.056	2
Lincoln	44.85	0.135	3.624	4.884	4
Mansfield	47.82	0.561	3.776	4.909	-2
Merthyr	44.92	0.412	3.528	4.377	-2
Newport	46.76	0.289	3.811	5.122	1
Northampton	49.24	0.296	3.499	4.925	1
Norwich	46.78	0.068	3.426	4.370	1
Nottingham	50.65	0.301	3.414	4.790	1
Nuneaton	50.09	0.684	3.712	4.930	-2
Oxford	44.95	0.009	3.958	5.524	4
Peterborough	45.62	0.019	3.551	4.850	1
Plymouth	43.56	0.295	3.377	4.277	-2
Portsmouth	45.69	0.319	3.391	4.404	5
Preston	52.24	0.312	3.389	4.576	-2
Rugby	48.03	0.019	3.722	5.486	0
Scunthorpe	44.33	0.748	3.383	4.774	1
Sheffield	48.89	0.552	3.371	4.550	-3
Southampton	44.84	0.222	3.272	4.372	4
Southport	44.94	0.128	3.085	4.185	3
Stoke	52.93	0.244	3.278	4.449	1
Sunderland	43.13	0.435	3.147	4.257	0
Swansea	43.28	0.413	3.603	4.539	-1
Swindon	44.40	0.008	3.543	4.789	0
Torquay	40.73	0.063	3.388	4.124	40
Wakefield	43.32	0.413	3.430	4.599	0
Warrington	50.00	0.456	3.608	4.844	1
West H'pool	40.73	0.449	3.332	4.386	0
Widnes	43.07	0.109	3.176	4.178	-2
Wigan	48.77	0.340	3.285	4.460	-1
Worcester	58.16	0.283	3.705	4.973	3
Worthing	36.72	0.003	3.546	4.736	8
York	47.94	0.014	3.331	4.625	1

KEY

(1) = Total labour force as percentage of residential population, 1961
(2) = Index of concentration of employment in 'old' industries (see text), 1961
(3) = Retail output per employee, 1961 (£)
(4) = Retail output per full-time employee, 1961 (£)
(5) = Index of hotel rating, 1961 (see text)

TABLE A7

Proxy Variables for Income: Counties, 1961

KEY: Foot of Table
UNITS: see key

	(1)	(2)	(3)	(4)	(5)	(6)
Cambridge	0.059	0.367	0.044	0.509	22.07	37.82
Cheshire	0.024	0.241	0.022	0.316	17.87	33.19
Cumberland	0.020	0.275	0.012	0.333	19.26	26.93
Devon	0.035	0.206	0.044	0.315	18.71	36.44
Durham	0.016	0.242	0.009	0.292	16.78	24.79
Glamorgan	0.037	0.233	0.023	0.320	18.04	28.16
Gloucester	0.027	0.322	0.031	0.413	20.56	34.15
Hampshire	0.029	0.250	0.040	0.355	21.69	41.74
Lancashire	0.024	0.297	0.019	0.380	22.24	30.26
Norfolk	0.034	0.227	0.018	0.319	16.92	27.87
Northampton	0.027	0.284	0.026	0.375	19.24	31.63
Nottingham	0.025	0.318	0.019	0.394	21.26	33.23
Oxford	0.031	0.343	0.048	0.454	24.02	39.53
Somerset	0.039	0.217	0.041	0.328	16.24	31.67
Stafford	0.022	0.316	0.016	0.392	21.61	28.33
Suffolk	0.033	0.235	0.034	0.329	16.66	28.43
Sussex	0.029	0.211	0.069	0.340	18.89	51.29
Warwick	0.022	0.405	0.022	0.488	23.70	32.59
Wiltshire	0.029	0.270	0.030	0.361	18.51	40.54
Worcester	0.025	0.280	0.026	0.364	20.66	31.69
Yorkshire E.	0.030	0.238	0.019	0.315	15.56	25.46
Yorkshire W.	0.024	0.300	0.021	0.382	17.70	26.55

KEY

(1) = Income from profits (£000's per capita)
(2) = Income from employment (£000's per capita)
(3) = Income from investment (£000's per capita)
(4) = Total income (£000's per capita)
(5) = Total non-domestic rateable value (per capita) (£)
(6) = Total non-industrial rateable value (per capita) (")

TABLE A8

Proxy Variables for Income: Counties

KEY: Foot of Table

UNITS; see key

	(1)	(2)	(3)	(4)	(5)
Cambridge	0.149	0.028	0.209	0.306	5946
Cheshire	0.135	0.024	0.186	0.288	4513
Cumberland	0.115	0.016	0.148	0.311	4840
Devon	0.141	0.024	0.192	0.295	6523
Durham	0.196	0.008	0.114	0.306	3849
Glamorgan	0.122	0.018	0.160	0.300	4897
Gloucester	0.143	0.024	0.194	0.314	5213
Hampshire	0.144	0.024	0.195	0.314	6361
Lancashire	0.104	0.013	0.129	0.301	4648
Norfolk	0.147	0.028	0.210	0.291	5307
Northampton	0.137	0.024	0.188	0.294	3898
Nottingham	0.125	0.017	0.162	0.298	4237
Oxford	0.145	0.029	0.207	0.331	5557
Somerset	0.151	0.029	0.214	0.299	5352
Stafford	0.126	0.018	0.166	0.312	3380
Suffolk	0.145	0.026	0.200	0.303	5133
Sussex	0.142	0.027	0.197	0.297	5963
Warwick	0.128	0.023	0.179	0.324	3763
Wiltshire	0.143	0.033	0.194	0.313	5661
Worcester	0.144	0.029	0.207	0.308	4265
Yorkshire E.	0.112	0.017	0.148	0.304	5378
Yorkshire W.	0.110	0.013	0.138	0.301	4150

KEY

(1) = Number of households with one car per capita
(2) = Number of households with two cars per capita
(3) = Number of cars per capita
(4) = Index of earnings in secondary sector (see text)
(5) = Percentage size of tertiary sector (index of)

APPENDIX B

THE RESULTS

TABLE B1

Income at County Level - $\bar{R}^2$

Independent Variable	Dependent Variable			
	Employment		Total	
	Normal	Log	Normal	Log
RATEABLE VALUE				
Domestic	-0.02	-0.01	-0.04	0.01
Shops	-0.01	-0.04	0.12	0.11
Offices	0.28	0.17	0.37	0.33
Other Commercial	0.03	0.01	0.07	0.09
Industrial	0.13	0.12	-0.01	-0.01
Crown	-0.04	0.03	-0.03	-0.02
Others	0.06	0.05	0.19	0.14
Non-Domestic	0.56	0.48	0.03	0.59
Non-Industrial	-0.04	-0.01	0.02	-0.04
Total	-0.03	0.02	0.03	0.02
CAR OWNERSHIP				
No. of households per population				
(1) one car	-0.01	0.03	-0.02	-0.01
(2) two cars	-0.04	0.01	0.02	0.05
No. of cars per population ..	-0.04	-0.02	0.01	0.01
INDEX OF EARNINGS	0.28	0.26	0.30	0.29

TABLE B2

Electricity Consumption and Number of Television Licences in Force as Proxy Variables for Income Against the Total Percentage of Employment in the Tertiary Sector 1961 as the Dependent Variable - $\bar{R}^2$

Independent Variable	Size of Tertiary Sector	
	Normal	Log
Electricity Consumption	0.06	0.08
Television Licences in Force	0.01	-0.02

TABLE B3

$\bar{R}^2$ Associated with Different Weighting Factors for the Population and Geographical Variables and the Total Percentage Employment in Tertiary Sector 1961 as the Dependent Variable

Weighting Factor	$\bar{R}^2$		
	Population	Large Town	Small Town
0	-0.01	-	-
1.0	-0.01	0.20	0.04
1.5	-0.04	0.17	0.05
2.0	-0.01	0.13	0.06
2.5	-0.01	0.09	0.06
3.0	-0.01	0.08	0.06
3.5	-0.01	0.07	0.06
4.0	-0.01	0.06	0.07

TABLE B4

Single and Multiregression Analysis with Original Independent Variables

Dependent Variable - percentage size of tertiary sector 1961

Sample - full sample

Contents - $\bar{R}^2$

Independent Variable		Results	
Name	No.	$\bar{R}^2$	Sign of Coefficient
Population	1	-0.05	positive
Large Town	2	0.20	negative
Small Town	3	0.04	positive
Export Sector	4	0.32	negative
Non-Domestic Rateable Value	5	0.01	positive
Index of Earnings	6	0.45	negative
Multiregression			
2 + 6		0.51	-
2 + 6 + 4		0.62	-
2 + 6 + 4 + 1		0.61	-
2 + 6 + 4 + 1 + 5		0.62	-

Significance of coefficients and commentary on residuals - see text

TABLE B5

Results for Single and Multiregression Analysis with Original Independent Variables

Dependent Variable - percentage size of tertiary sector

Sample - 38 and 55

Contents - $\bar{R}^2$

Independent Variable		Results			
		Sample 38		Sample 55	
Name	No.	$\bar{R}^2$	Coefficient	$\bar{R}^2$	Coefficient
Population (log)	1	0.48	+	-0.01	+
Large Town	2	0.08	-	0.31	-
Small Town	3	-0.03	+	0.01	+
Export Sector	4	0.47	-	0.29	-
Non-Domestic Rateable Value	5	-0.02	-	0.03	-
Index of Earnings	6	0.18	-	0.19	-
Multiregression					
1 + 2		0.47		0.30	
1 + 2 + 6		0.50		0.29	
1 + 2 + 6 + 4		0.60		0.54	
1 + 2 + 6 + 4 + 5		0.60		0.58	

For significance of coefficients - see text

TABLE B6

Results of Single and Multiregression Analysis with Original Independent Variables

Dependent Variables - percentage size of individual industries

Sample - full sample

Contents - $\bar{R}^2$

Independent Variable		Industrial Order						
Name	No.	18	19	20	21	22	23	24
Large Town	1	0.01	0.02	0.13	0.08	0.08	0.07	0.02
Index of Earnings	2	0.07	0.06	0.28	0.17	0.17	0.03	0.05
Export Sector	3	0.05	-0.02	0.33	0.07	0.07	0.22	0.10
Multiregression								
1 + 2		-0.01	0.02	0.35	0.23	0.22	0.09	0.05
1 + 2 + 3		0.04	0.01	0.50	0.23	0.23	0.22	0.11

KEY: Industrial Orders

18 = Public Utilities

19 = Transport

20 = Distributive Trade

21 = Finance, banking and Insurance

22 = Professional and Scientific

23 = Miscellaneous

24 = Public Administration

TABLE B7

Single and Multiregression Analysis
with Original Independent Variables

Dependent Variable - percentage size of individual industries successively subtracted from total

Sample - full sample
Contents - $\bar{R}^2$

Independent Variable		Industrial Orders Deleted						
Name	No.	Total	18	18, 19	18, 19 22	18, 19 22, 24	18, 19, 22 24, 21	18, 19, 22 24, 21, 23
Large Town	1	0.20	0.01	-0.02	-0.01	-0.01	0.01	0.01
Index of Earnings	2	0.45	0.46	0.49	0.40	0.34	0.29	0.36
Exports	3	0.32	0.31	0.25	0.23	0.16	0.16	0.31
Multiregression								
1 + 2		0.51	0.26	0.21	0.16	0.11	0.09	0.15
1 + 2 + 3		0.62	0.52	0.45	0.37	0.24	0.21	0.43

KEY: See foot of Table B6

TABLE B8

Single and Multiregression Analysis
with Original Independent Variables

Dependent Variable - percentage size of tertiary sector
Sample - North/south
Contents - $\bar{R}^2$

Independent Variable		Sample	
Name	No.	North	South
Large Town	1	-0.03	0.39
Index of Earnings	2	0.15	0.60
Exporting Sector	3	0.30	0.32
Multiregression			
1 + 2		0.18	0.74
1 + 2 + 3		0.47	0.75

TABLE B9

Single Regression Analysis with Log Transformation of Original Independent Variables

Dependent Variable - percentage size of tertiary sector

Sample - full sample

Contents - $\bar{R}^2$

Independent Variable (log version)	Dependent Variable	
	Normal	Log
Population	0.04	-0.01
Large Town	0.06	0.07
Small Town	-0.05	-0.05
Export Sector	0.12	0.12
Non-Domestic Rateable Value	0.01	-0.05
Index of Earnings	0.15	0.13

TABLE B10

Single and Multiregression Analysis with Original Independent Variables in the First Differences Version of the Equilibrium Model

Dependent Variable - change in the percentage employment in tertiary sector 1951 - 1961

Sample - full sample

Contents - $\bar{R}^2$ and coefficients

Independent Variable		Results	
Name	No.	$\bar{R}^2$	Coefficient
Change in Population	1	-0.07	positive
Change in Large Town	2	-0.01	negative
Change in Small Town	3	-0.01	negative
Change in Index of Earnings	4	0.07	positive
Change in Export Sector	5	-0.02	negative
Change in Non-Domestic Rateable Value	6	-0.04	negative
Multiregression			
4 + 2		0.07	
4 + 2 + 1		0.07	
4 + 2 + 1 + 3		0.07	
4 + 2 + 1 + 3 + 5		0.06	
4 + 2 + 1 + 3 + 5 + 6		0.06	

TABLE B11

Single Regression Analysis for Log Transformations of Original Independent Variables in First Differences Version of the Equilibrium Model

Dependent Variable - change in percentage employment in tertiary sector 1951 - 1961

Sample - full sample

Contents - $\bar{R}^2$

Independent Variable (Log)	Dependent Variable	
	Normal	Log
Change in Population	-0.07	-0.07
Change in Large Town	-0.01	-0.02
Change in Small Town	-0.01	-0.03
Change in Index of Earnings	0.07	0.06
Change in Export Sector	-0.02	-0.02
Change in Non-Domestic Rateable Value	-0.04	-0.05

TABLE B12

Single and Multiregression Analysis with Original Independent Variables in the First Differences Version of the Equilibrium Model

Dependent Variable - change in percentage employment in tertiary sector 1951 - 1961

Sample - 38 and 55

Contents - $\bar{R}^2$ and coefficients

Independent Variable		Sample 38		Sample 55	
Name	No.	$\bar{R}^2$	Coefficients	$\bar{R}^2$	Coefficients
Change in Population	1	-0.05	positive	-0.05	positive
Change in Large Town	2	-0.04	negative	-0.03	negative
Change in Small Town	3	-0.01	negative	-0.02	negative
Change in Index of Earnings	4	0.02	positive	0.04	positive
Change in Export Sector	5	-0.05	negative	-0.05	negative
Change in Non-Domestic Rateable Value	6	0.01	negative	0.01	negative
Multiregression					
4 + 2		0.02		0.04	
4 + 2 + 1		0.02		0.04	
4 + 2 + 1 + 3		0.02		0.03	
4 + 2 + 1 + 3 + 5		0.01		0.01	
4 + 2 + 1 + 3 + 5 + 6		0.01		0.01	

TABLE B13

Single and Multiregression Analysis with Original Independent Variables in the First Differences Version of the Equilibrium Model

Dependent Variable - change in the percentage employment in the tertiary sector 1951 - 1961

Sample - North/South

Contents - $\bar{R}^2$ and coefficients

Independent Variable		North		South	
Name	No.	$\bar{R}^2$	Coefficient	$\bar{R}^2$	Coefficient
Change in Population	1	-0.06	positive	-0.02	positive
Change in Large Town	2	0.05[1]	positive	-0.02	positive
Change in Small Town	3	-0.01	positive	-0.01	positive
Change in Index of Earnings	4	0.06	positive	0.01	positive
Change in Export Sector	5	-0.03	negative	-0.02	positive
Change in non-Domestic Rateable Value	6	-0.03	positive	-0.04	positive
Multiregression					
4 + 2		0.11		-0.01	
4 + 2 + 5		0.12		0.01	
4 + 2 + 5 + 3		0.11		0.02	
4 + 2 + 5 + 3 + 1		0.11		0.02	
4 + 2 + 5 + 3 + 1 + 6		0.11		0.02	

[1] Result not significant at 95% confidence level

TABLE B14

Single and Multiregression Analysis with the Original Independent Variables in the Disequilibrium Model Based on 1951 and 1961 Data

Dependent Variable - change in percentage employment in tertiary sector 1951 - 1961

Sample - full sample

Contents - $\bar{R}^2$

Independent Variable		1951 Data	1961 Data
Name	No.		
Population	1	-0.07	-0.04
Large Town	2	-0.01	-0.02
Small Town	3	0.03	-0.04
Export Sector	4	-0.02	-0.01
Non-Domestic Rateable Value	5	-0.01	-0.01
Index of Earnings	6	-0.24	-0.01
Multiregression			
2 + 6		0.22	-0.04
2 + 6 + 4		0.20	-0.04
2 + 6 + 4 + 1		0.21	0.01
2 + 6 + 4 + 1 + 3		0.20	-0.03
2 + 6 + 4 + 1 + 3 + 5		0.20	-0.02

TABLE B15

Single and Multiregression Analysis with Original Independent Variables in Disequilibrium Model Based on 1951 and 1961 Data

Dependent Variable - change in employment in tertiary sector 1951 to 1961

Sample - North/South
Contents - $\bar{R}^2$

Independent Variable		North		South	
Name	No.	1951 Data	1961 Data	1951 Data	1961 Data
Large Town	1	-0.01	0.01	-0.02	-0.01
Export Sector	2	-0.03	-0.03	0.03	-0.04
Index of Earning	3	0.21	0.02	-0.19	0.02
Multiregression					
1 + 3		0.21	0.03	0.19	-0.03
1 + 3 + 2		0.21	0.02	0.20	-0.03

TABLE B16

Single and Multiregression Analysis for Original Independent Variables in Disequilibrium Model Based on 1951 and 1961 Data

Dependent Variable - change in employment in tertiary sector 1951 and 1961

Sample - 38 and 55

Contents - $\bar{R}^2$

Independent Variable		Sample 38		Sample 55	
Name	No.	1951 Data	1961 Data	1951 Data	1961 Data
Population	1	0.03	-0.02	-0.01	-0.01
Large Town	2	0.05	0.02	-0.05	-0.01
Small Town	3	0.09	-0.01	0.06	0.03
Export Sector	4	-0.03	-0.02	-0.01	0.03
Non-Domestic Rateable Value	5	0.02	-0.01	0.02	0.01
Index of Earnings	6	0.20	-0.01	0.21	-0.01
Multiregression					
2 + 6		0.21	0.01	0.21	-0.02
2 + 6 + 4		0.21	-0.02	0.20	-0.03
2 + 6 + 4 + 1		0.20	-0.03	0.20	-0.05
2 + 6 + 4 + 1 + 3		0.21	-0.06	0.20	-0.05
2 + 6 + 4 + 1 + 3 + 5		0.21	-0.06	0.20	-0.06

TABLE B17

Single Regression Analysis with Original Independent Variables (Log Transformation) in Disequilibrium Model Based on 1951 and 1961 Data

Dependent Variable - percentage change in size of tertiary sector 1951 - 1961

Sample - full sample

Contents - $\bar{R}^2$

Independent Variable (log version)	1951 Data		1961 Data	
	Normal	Log	Normal	Log
Population	-0.05	-0.05	-0.01	-0.01
Large Town	-0.01	-0.01	-0.05	-0.06
Small Town	-0.04	-0.04	-0.02	0.02
Export Sector	-0.02	-0.02	-0.04	0.05
Non-Domestic Rateable Value	0.04	0.03	-0.03	-0.01
Index of Earnings	0.22	0.18	0.01	0.03

TABLE B18

Results for New Independent Variables Introduced as Extensions to the Original Model

Dependent Variable - percentage employment in tertiary sector 1961
Sample - full sample
Contents - $\bar{R}^2$ and sign of coefficient

Independent Variable	Results	
	$\bar{R}^2$	Sign of Coefficient
Index of Concentration	0.20	negative
Index of Hotel Accommodation	0.25	positive
Index of Commuting	0.17	negative
Output per employee*	0.05	negative
Output per full-time employee	0.03	negative
Population change from 1901	-0.02	positive
Population change from 1911	-0.01	negative
Population change from 1921	0.02	positive
Population change from 1931	0.01	positive
Population change from 1951	0.01	positive
Percentage Size of Tertiary Sector 1931 ‡	0.74	positive

* Note results only for distribution as dependent variable

‡ Results based on restricted sample - see text

TABLE B19

Results Obtained When the Income of the Hinterland is Considered with the Index of Earnings as Independent Variables for the Service Centres

Dependent Variable - percentage size of tertiary sector

Sample - service centres

Contents - $\bar{R}^2$

Independent Variable	$\bar{R}^2$
Index of Earnings	0.49
Index of Earnings + Index 1	0.45
Index of Earnings + Index 2	0.47
Index of Earnings + Index 3	0.40

KEY:

Index 1 = Index of hinterlands based on average index of earnings

Index 2 = Index of hinterlands income based on actual earnings

Index 3 = Index of hinterlands income based on actual earnings weighted by distance

TABLE B20

Single and Multiregression Analysis with Hypotheses Extending the Original Model as Independent Variables

Dependent Variable - percentage size of tertiary sector 1961
Sample - full sample
Contents - $\bar{R}^2$ and coefficients

Independent Variable		Results	
Name	No.	$\bar{R}^2$	Coefficient
Index of Concentration	1	0.20	negative
Index of Commuting	2	0.17	negative
Size of Tertiary Sector 1931	3	0.74	positive
Index of Hotel Accommodation	4	0.25	positive
Multiregression with Original Model (see footnote)			
Original Model		0.62	
Original Model + 1		0.66	
Original Model + 1 + 2		0.67	
Original Model + 1 + 2 + 3		0.82	
Original Model + 1 + 2 + 3 + 4		0.83	
Original Model + 3		0.82	
Original Model + 3 + 4		0.83	
Original Model + 3 + 4 + 1		0.83	
Original Model + 3 + 4 + 1 + 2		0.83	

Note - Original consists of large town, index of earnings and export sector variables

TABLE B21

Single and Multiregression Analysis with Hypotheses Extending Original Model as Independent Variables

Dependent Variable - percentage size of tertiary sector 1961

Sample - North and South

Contents - $\bar{R}^2$ and coefficients

Independent Variables		North		South	
Name	No.	$\bar{R}^2$	Coefficient	$\bar{R}^2$	Coefficient
Index of Concentration	1	0.27	negative	0.05	negative
Index of Commuting	2	0.04	negative	0.23	negative
Percentage size of tertiary sector 1931	3	0.69	positive	0.74	positive
Index of Hotel Accommodation	4	0.18	positive	0.23	positive
Multiregression with Original Model (see footnote)					
Original model		0.47		0.75	
Original model + 1		0.52		0.76	
Original model + 1 + 2		0.60		0.76	
Original model + 1 + 2 + 3		0.81		0.80	
Original model + 1 + 2 + 3 + 4		0.84		0.80	
Original model + 3		0.78		0.80	
Original model + 3 + 4		0.81		0.80	
Original model + 3 + 4 + 1		0.81		0.80	
Original model + 3 + 4 + 1 + 2		0.84		0.80	

Note - Original model consists of the large town, index of earnings and export sector variables

TABLE B22

Single and Multiregression Analysis with Log Transformations of Hypotheses Extending the Original Model as Independent Variables

Dependent Variable: Percentage size of tertiary sector 1961

Sample: Full sample

Contents: $\bar{R}^2$

Independent Variable (log)		Dependent Variable	
Name	No.	Normal	Log
Index of Concentration	1	0.18	0.17
Index of Commuting	2	0.17	0.17
Percentage Size of Tertiary Sector 1931	3	0.72	0.73
Index of Hotel Accommodation	4	0.24	0.23
Multiregression With Original Model (see footnote)			
Original Model		0.62	0.61
Original Model + 1		0.65	0.63
Original Model + 1 + 2		0.65	0.64
Original Model + 1 + 2 + 3		0.82	0.78
Original Model + 1 + 2 + 3 + 4		0.83	0.80
Original Model + 3		0.82	0.78
Original Model + 3 + 4		0.83	0.80
Original Model + 3 + 4 + 1		0.83	0.80
Original Model + 3 + 4 + 1 + 2		0.83	0.80

Note - Original model consists of large town, index of earnings and export sector variables.

TABLE B23

Single and Multiregression Analysis with Hypotheses Extending the Original First Differences Version of the Equilibrium Model

Dependent Variable - the change in percentage size of tertiary sector 1951 - 1961

Sample - full sample

Contents - $\bar{R}^2$

Independent Variable		Results	
Name	No.	$\bar{R}^2$	Coefficient
Index of Concentration	1	0.08	positive
Index of Commuting	2	0.02	positive
Size of Tertiary Sector 1931	3	0.12	positive
Index of Hotel Accommodation	4	0.01	positive
Multiregression with Original Model (see footnote)			
Original Model		0.07	
Original Model + 1		0.07	
Original Model + 1 + 2		0.07	
Original Model + 1 + 2 + 3		0.11	
Original Model + 1 + 2 + 3 + 4		0.10	
Original Model + 3		0.11	
Original Model + 3 + 4		0.10	
Original Model + 3 + 4 + 1		0.10	
Original Model + 3 + 4 + 1 + 2		0.10	

Note - Original model consists of the change in the index of earnings 1951 - 1961

TABLE B24

Single and Multiregression Analysis with Hypotheses Extending the Original First Differences Version of the Equilibrium Model

Dependent Variable - change in percentage size of tertiary sector 1951 - 1961

Sample - North/South

Contents - $\bar{R}^2$ and coefficients

Independent Variable		North		South	
Name	No.	$\bar{R}^2$	Coefficient	$\bar{R}^2$	Coefficient
Index of Concentration	1	-0.02	negative	-0.04	negative
Index of Commuting	2	-0.01	negative	-0.01	negative
Size of tertiary sector 1931	3	-0.07	positive	-0.06	positive
Index of hotel accommodation	4	0.01	positive	-0.01	positive
Multiregression with Original Model (see footnote)					
Original Model		0.12		0.02	
Original Model + 1		0.12		-0.02	
Original Model + 1 + 2		0.11		-0.06	
Original Model + 1 + 2 + 3		0.11		-0.11	
Original Model + 1 + 2 + 3 + 4		0.11		-0.16	
Original Model + 3		0.11		-0.01	
Original Model + 3 + 4		0.11		-0.05	
Original Model + 3 + 4 + 1		0.11		-0.08	
Original Model + 3 + 4 + 1 + 2		0.12		-0.16	

Note - Original model consists of the change in the index of earnings 1951 - 1961

TABLE B25

Single Regression Analysis with Log Transformation
of Hypotheses Extending Original First Differences
Version of Equilibrium Model

Dependent Variable - change in percentage size of tertiary sector 1951 - 1961

Sample - full sample
Contents - $\bar{R}^2$

Independent Variable (log)	Dependent Variable	
	Normal	Log
Index of Concentration	-0.01	-0.01
Index of Commuting	-0.07	-0.05
Size of Tertiary Sector 1931	-0.06	-0.08
Index of Hotel Accommodation	-0.01	-0.01

TABLE B26

Single and Multiregression Analysis with the New Hypothesis Extending the Original Model as Independent Variables in the Disequilibrium Model

Dependent Variable - change in the percentage size of tertiary sector 1951 - 1961

Sample - full sample

Contents - $\bar{R}^2$ and coefficients

Independent Variable: Name	No.	$\bar{R}^2$	Coefficient
Index of Concentration	1	0.08	positive
Index of Commuting	2	0.02	positive
Percentage Size of Tertiary Sector 1931	3	0.12	positive
Index of Hotel Accommodation	4	0.01	positive
Multiregression With Original Model (see footnote)			
Original Model		0.01	
Original Model + 1		0.04	
Original Model + 1 + 2		0.04	
Original Model + 1 + 2 + 3		0.08	
Original Model + 1 + 2 + 3 + 4		0.08	
Original Model + 3		0.08	
Original Model + 3 + 4		0.08	
Original Model + 3 + 4 + 1		0.08	
Original Model + 3 + 4 + 1 + 2		0.08	

Note - Original model consists of large town, index of earnings and export variables

TABLE B27

Single and Multiregression Analysis with the New Hypotheses Extending the Original Model as Independent Variables in the Disequilibrium Model

Dependent Variable - change in the percentage size of the tertiary sector 1951 - 1961

Sample - north/south

Contents - $\bar{R}^2$ and coefficients

Independent Variable		North		South	
Name	No.	$\bar{R}^2$	Coefficient	$\bar{R}^2$	Coefficient
Index of Concentration	1	-0.01	negative	-0.01	negative
Index of Commuting	2	-0.02	negative	-0.04	negative
Size of tertiary sector 1931	3	0.03	positive	0.01	positive
Index of Hotel Accommodation	4	-0.03	positive	-0.03	positive
Multiregression With Original Model (see footnote)					
Original Model		-0.13		-0.05	
Original Model + 1		-0.15		-0.09	
Original Model + 1 + 2		-0.16		-0.14	
Original Model + 1 + 2 + 3		-0.18		-0.14	
Original Model + 1 + 2 + 3 + 4		-0.24		-0.14	
Original Model + 3		-0.23		-0.14	
Original Model + 3 + 4		-0.24		-0.14	
Original Model + 3 + 4 + 1		-0.24		-0.14	
Original Model + 3 + 4 + 1 + 2		-0.24		-0.14	

Note - The original model consists of large town, index of earnings and export sector variables

TABLE B28

Single Regression Analysis With Log Transformations of New Hypotheses Extending the Original Model as Independent Variables in the Disequilibrium Model

Dependent Variable - change in percentage size of tertiary sector 1951 - 1961

Sample - full sample
Contents - $\bar{R}^2$

Independent Variable (log)	Dependent Variable	
	Normal	Log
Index of Concentration	-0.03	-0.07
Index of Commuting	0.01	-0.03
Size of Tertiary Sector 1931	0.03	0.02
Index of Hotel Accommodation	-0.03	-0.04

TABLE B29

Single and Multiregression Analysis with Original and New Hypotheses in the First Equilibrium Version of the Model Based Upon the Absolute Size of the Tertiary Sectors

Dependent Variable - numbers employed in tertiary sector 1961

Sample - full sample

Contents - $\bar{R}^2$ and coefficients

Independent Variable		$\bar{R}^2$	Coefficients
Name	No.		
Population (log)[1]	1	0.70	positive
Large Town	2	0.09	negative
Export Sector	3	-0.01	negative
Index of Earnings	4	-0.01	positive
Index of Concentration	5	-0.02	negative
Index of Commuting	6	-0.02	negative
Size of Tertiary Sector 1931	7	0.03	positive
Index of Hotel Accommodation	8	-0.01	positive
Multiregression Analysis			
1 + 2		0.71	
1 + 2 + 3		0.71	
1 + 2 + 3 + 4		0.71	
1 + 2 + 3 + 4 + 5		0.71	
1 + 2 + 3 + 4 + 5 + 6		0.72	
1 + 2 + 3 + 4 + 5 + 6 + 7		0.72	
1 + 2 + 3 + 4 + 5 + 6 + 7 + 8		0.77	

[1] The normal version had $\bar{R}^2 = 0.64$ (see text)

TABLE B30

Single and Multiregression Analysis of Original and New Hypotheses in the First Equilibrium Version of the Model Based Upon the Absolute Size of the Tertiary Sector

Dependent Variable - absolute size of tertiary sector

Sample - full sample

Contents - $\bar{R}^2$ and coefficients

Independent Variable		$\bar{R}^2$	Coefficients
Name	No.		
Labour Force	1	0.58	positive
Large Town	2	0.09	negative
Export Sector	3	-0.01	negative
Index of Earnings	4	-0.01	positive
Index of Concentration	5	-0.02	negative
Index of Commuting	6	-0.02	negative
Size of Tertiary Sector 1931	7	0.03	positive
Index of Hotel Accommodation	8	-0.01	negative
Multiregression Analysis			
1 + 2		0.59	
1 + 2 + 3		0.64	
1 + 2 + 3 + 4		0.64	
1 + 2 + 3 + 4 + 5		0.64	
1 + 2 + 3 + 4 + 5 + 6		0.65	
1 + 2 + 3 + 4 + 5 + 6 + 7		0.66	
1 + 2 + 3 + 4 + 5 + 6 + 7 + 8		0.72	

TABLE B31

Multiregression Analysis with the Original and New Hypotheses of the Equilibrium Version of the Model Based upon the Absolute Size of the Tertiary Sector 1961

Dependent Variable - absolute size of the tertiary sector 1961

Sample - full sample

Contents - $\bar{R}^2$

Independent Variables (log) (see footnote)	Absolute Size of Tertiary Sector	
	Normal	Log
Size of labour force	0.89	0.70
Labour force + 2	0.89	0.70
Labour force + 2 + 3	0.89	0.71
Labour force + 2 + 3 + 4	0.89	0.71
Labour force + 2 + 3 + 4 + 5	0.89	0.71
Labour force + 2 + 3 + 4 + 5 + 6	0.89	0.75
Labour force + 2 + 3 + 4 + 5 + 6 + 7	0.89	0.75
Labour force + 2 + 3 + 4 + 5 + 6 + 7 + 8	0.89	0.80

Note - The coding for the independent variables is the same as in Table B30

TABLE B32

Single and Multiregression Analysis with Original and New Hypotheses in First Differences Version of Equilibrium Model Based Upon the Absolute Size of the Tertiary Sector 1961

Dependent Variable - change in absolute size of tertiary sector

Sample - full sample

Contents - $\bar{R}^2$ and coefficients

Independent Variable		$\bar{R}^2$	Coefficients
Name	No.		
Change in Labour Force	1	-0.57	positive
Change in Large Town	2	-0.01	positive
Change in Export Sector	3	-0.02	positive
Change in Index of Earnings	4	-0.03	negative
Index of Concentration	5	-0.02	positive
Index of Commuting	6	-0.01	negative
Size of Tertiary Sector 1931	7	-0.01	negative
Index of Hotel Accommodation	8	-0.02	positive
Multiregression Analysis			
1 + 2		0.57	
1 + 2 + 3		0.57	
1 + 2 + 3 + 4		0.56	
1 + 2 + 3 + 4 + 5		0.57	
1 + 2 + 3 + 4 + 5 + 6		0.57	
1 + 2 + 3 + 4 + 5 + 6 + 7		0.57	
1 + 2 + 3 + 4 + 5 + 6 + 7 + 8		0.57	

TABLE B33

Single and Multiregression Analysis with Original and New Hypotheses in Disequilibrium Model Based Upon the Absolute Size of the Tertiary Sector

Dependent Variable - change in absolute size of the tertiary sector

Sample - full sample

Contents - $\bar{R}^2$ and coefficients

Independent Variable: Name	No.	$\bar{R}^2$	Coefficients
Size of Labour Force	1	-0.02	positive
Large Town	2	-0.01	negative
Export Sector	3	0.14	negative
Index of Earnings	4	-0.01	negative
Index of Concentration	5	-0.02	positive
Index of Commuting	6	-0.01	negative
Size of Tertiary Sector 1931	7	-0.01	negative
Index of Hotel Accommodation	8	-0.02	positive
Multiregression Analysis			
1 + 2		-0.01	
1 + 2 + 3		0.02	
1 + 2 + 3 + 4		0.11	
1 + 2 + 3 + 4 + 5		0.09	
1 + 2 + 3 + 4 + 5 + 6		0.07	
1 + 2 + 3 + 4 + 5 + 6 + 7		0.05	
1 + 2 + 3 + 4 + 5 + 6 + 7 + 8		0.03	

TABLE B34

Single and Multiregression Analysis with Original and New Hypotheses in Disequilibrium Model Based Upon the Absolute Size of the Tertiary Sector

Dependent Variable - change in absolute size of tertiary sector
Sample - full sample
Contents - $\bar{R}^2$

Independent Variable (log)		Dependent Variable	
Name	No.	Normal	Log
Size of Labour Force	1	-0.01	-0.02
Large Town	2	-0.03	-0.01
Export Sector	3	0.05	0.12
Index of Earnings	4	-0.02	-0.02
Index of Concentration	5	0.01	-0.01
Index of Commuting	6	-0.03	-0.01
Size of Tertiary Sector 1931	7	-0.03	-0.01
Index of Hotel Accommodation	8	-0.04	-0.01
Multiregression Analysis			
1 + 2		-0.01	-0.01
1 + 2 + 3		-0.01	0.02
1 + 2 + 3 + 4		0.03	0.09
1 + 2 + 3 + 4 + 5		0.01	0.05
1 + 2 + 3 + 4 + 5 + 6		-0.01	0.01
1 + 2 + 3 + 4 + 5 + 6 + 7		-0.01	0.01
1 + 2 + 3 + 4 + 5 + 6 + 7 + 8		-0.01	0.01

TABLE B35

Single and Multiregression Analysis with Original and New Hypotheses in Equilibrium Version of Model Based Upon the Absolute Size of the Tertiary Sector

Dependent Variable - numbers employed in tertiary sector 1961

Sample - North/South

Contents - $\bar{R}^2$ and coefficients

Independent Variable		North		South	
Name	No.	$\bar{R}^2$	Coefficients	$\bar{R}^2$	Coefficients
Population (Log)	1	0.25	positive	0.78	positive
Large Town	2	0.04	negative	0.08	negative
Export Sector	3	-0.03	positive	-0.01	positive
Index of Earnings	4	-0.03	negative	-0.03	negative
Index of Concentration	5	-0.03	negative	-0.01	negative
Index of Commuting	6	-0.01	negative	-0.04	negative
Size of Tertiary Sector 1931	7	-0.02	negative	0.01	negative
Index of Hotel Accommodation	8	-0.01	positive	-0.03	positive
Multiregression Analysis					
1 + 2		0.27		0.81	
1 + 2 + 3		0.26		0.82	
1 + 2 + 3 + 4		0.21		0.82	
1 + 2 + 3 + 4 + 5		0.18		0.81	
1 + 2 + 3 + 4 + 5 + 6		0.14		0.80	
1 + 2 + 3 + 4 + 5 + 6 + 7		0.11		0.80	
1 + 2 + 3 + 4 + 5 + 6 + 7 + 8		0.50		0.79	

TABLE B36

Single and Multiregression Analysis with Original and New Hypotheses in the Equilibrium Version of Model Based Upon the Absolute Size of the Tertiary Sector

Dependent Variable - absolute size of tertiary sector

Sample - full sample

Contents - $\bar{R}^2$

Independent Variable (log)		Absolute Size of Tertiary Sector	
Name	No.	Normal	Log
Population	1	0.70	0.63
Large Town	2	0.04	0.04
Export Sector	3	-0.01	-0.03
Index of Earnings	4	-0.02	0.01
Index of Concentration	5	-0.01	-0.03
Index of Commuting	6	-0.03	0.01
Size of Tertiary Sector 1931	7	0.01	0.03
Index of Hotel Accommodation	8	-0.01	-0.01
Multiregression Analysis			
1 + 2		0.70	0.63
1 + 2 + 3		0.70	0.65
1 + 2 + 3 + 4		0.71	0.66
1 + 2 + 3 + 4 + 5		0.71	0.66
1 + 2 + 3 + 4 + 5 + 6		0.71	0.66
1 + 2 + 3 + 4 + 5 + 6 + 7		0.72	0.66
1 + 2 + 3 + 4 + 5 + 6 + 7 + 8		0.71	0.66

APPENDIX C

SCATTER DIAGRAMS

diagram C1

Total Non-Domestic Rateable Value against Total Income

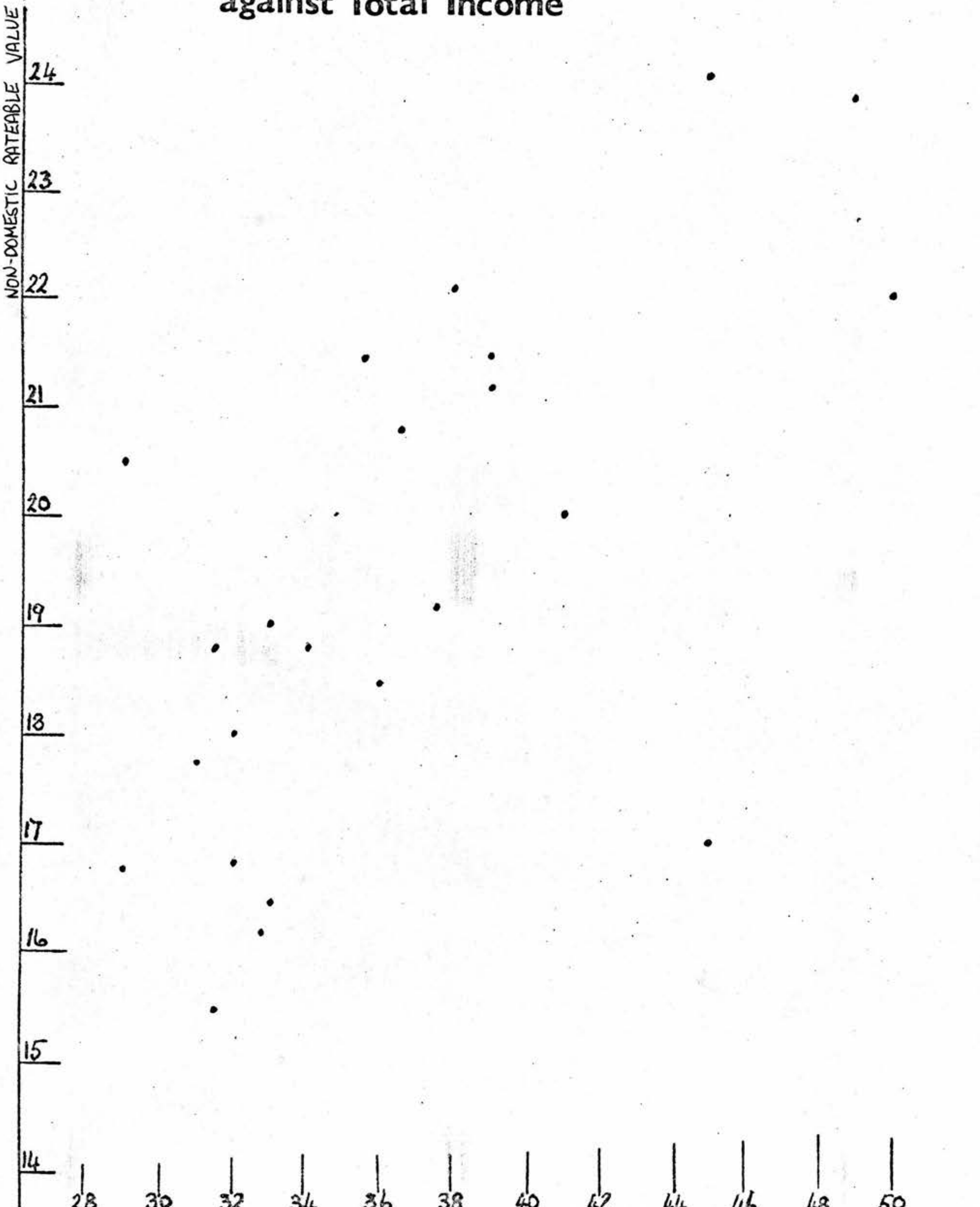


diagram C2

Index of Earnings against Total Income

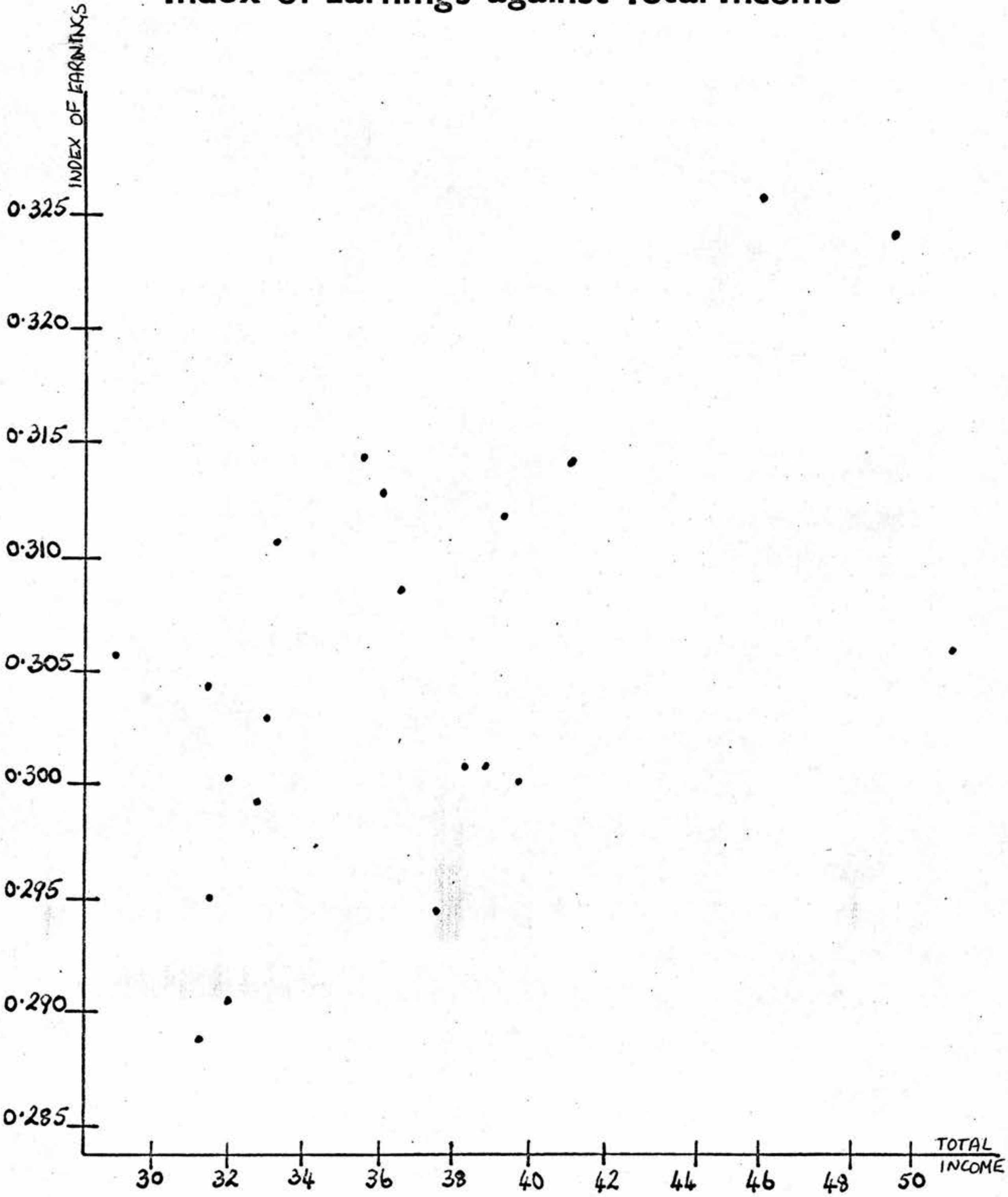
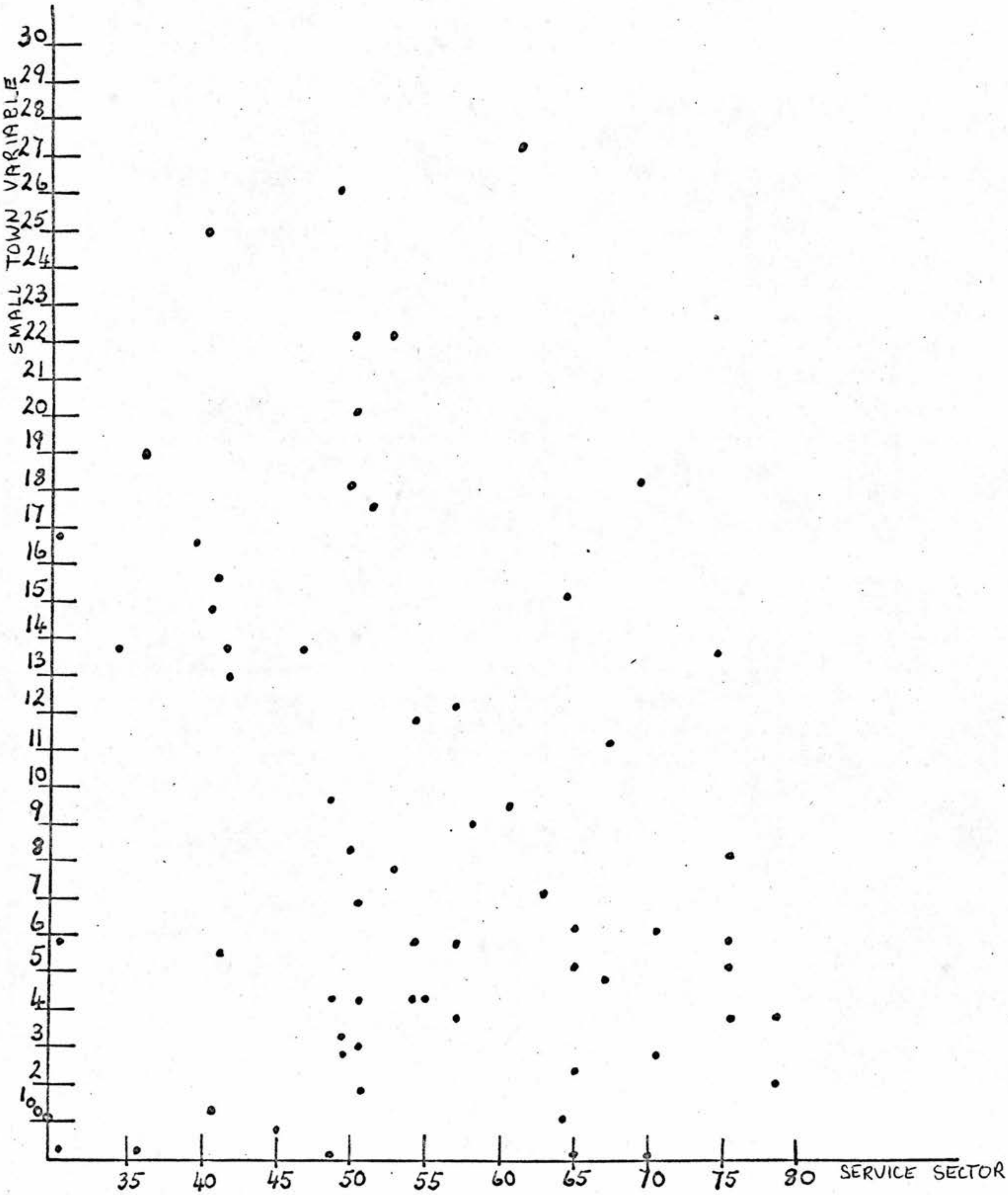


diagram C3

Size of Small Town Variable against Service Sector



Non–Domestic Rateable Value against Size of Tertiary Sector

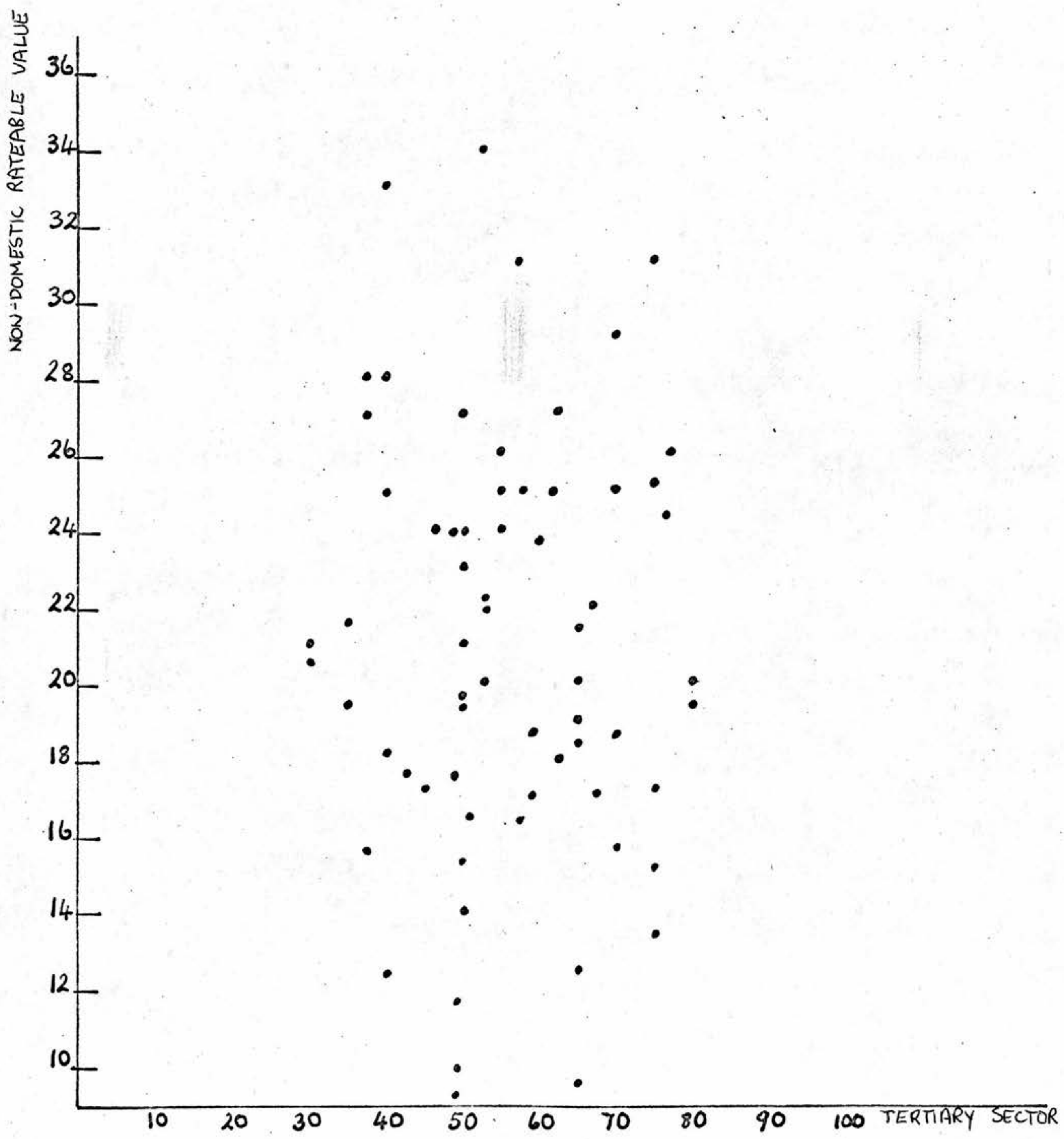


diagram C5

Population against Tertiary Sector

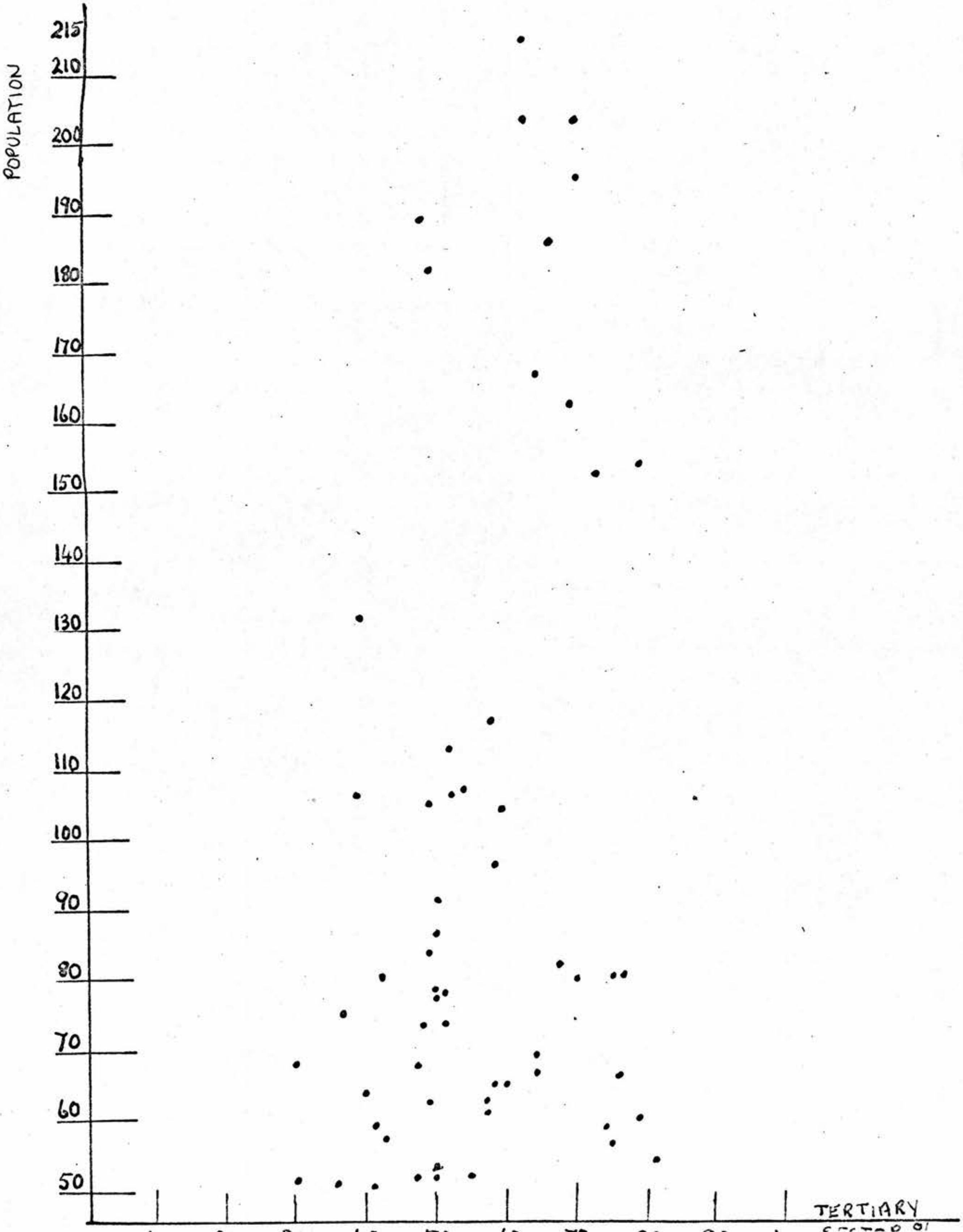


diagram C6

Large Town Variable against Size of Tertiary Sector

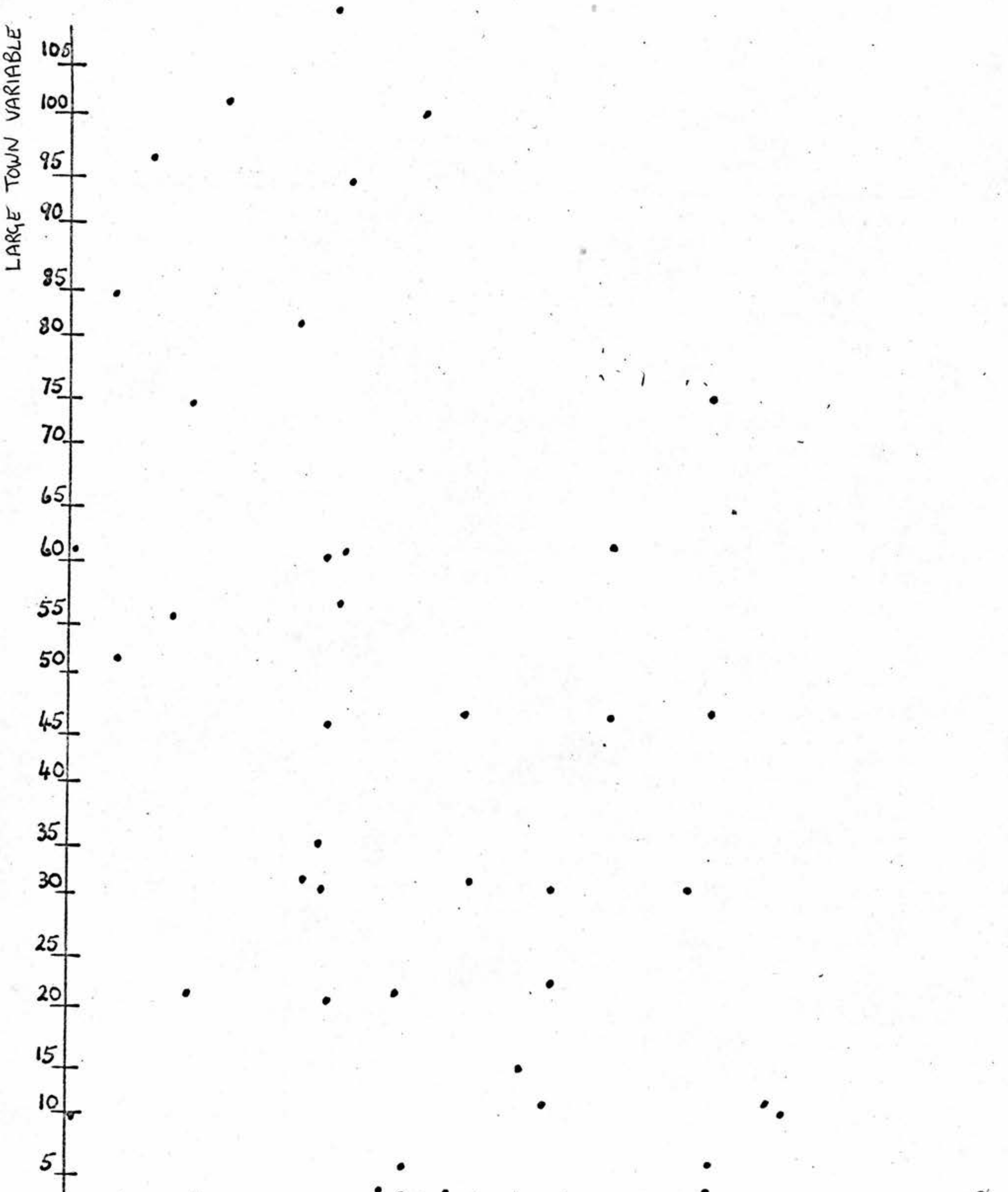
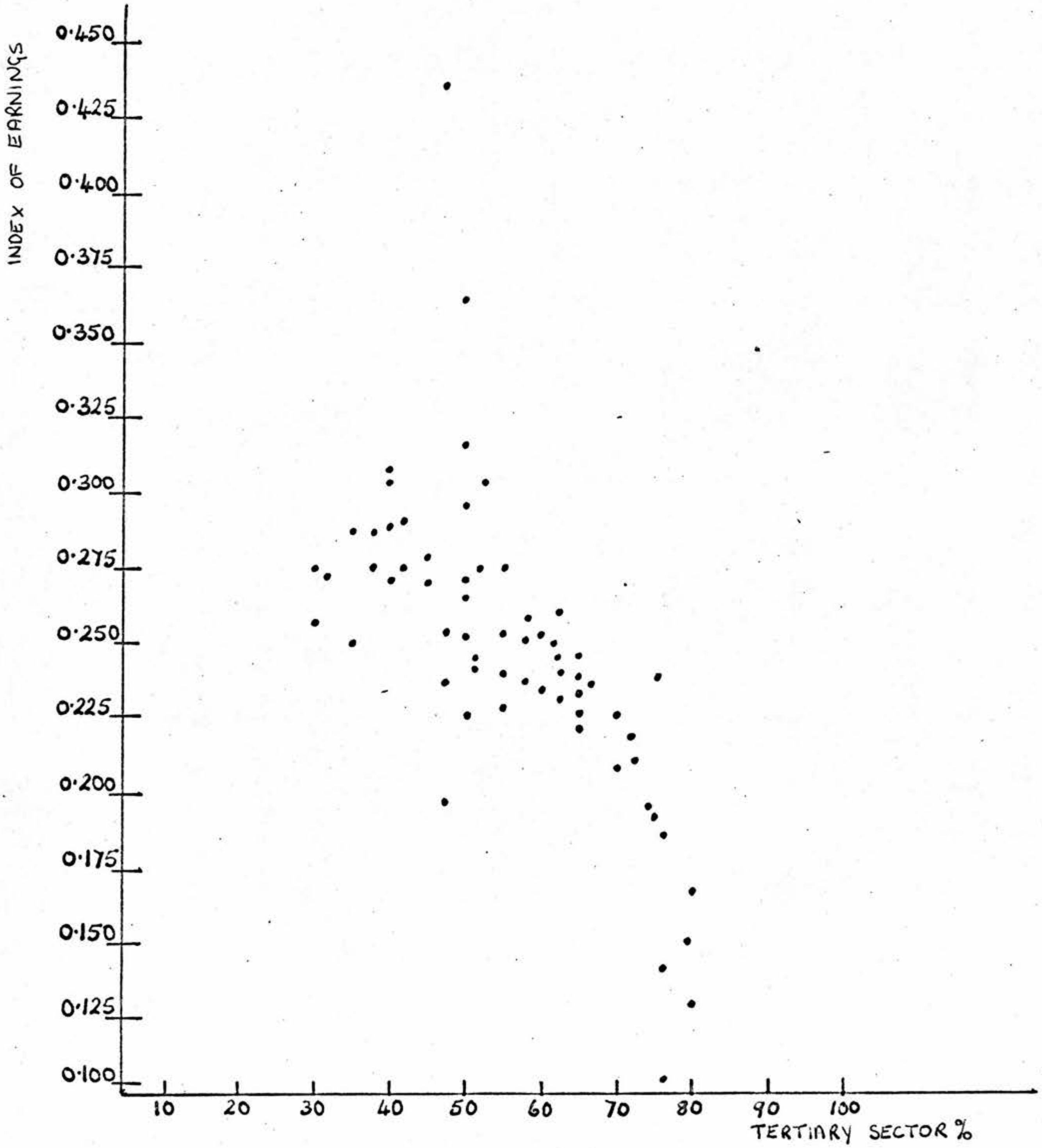


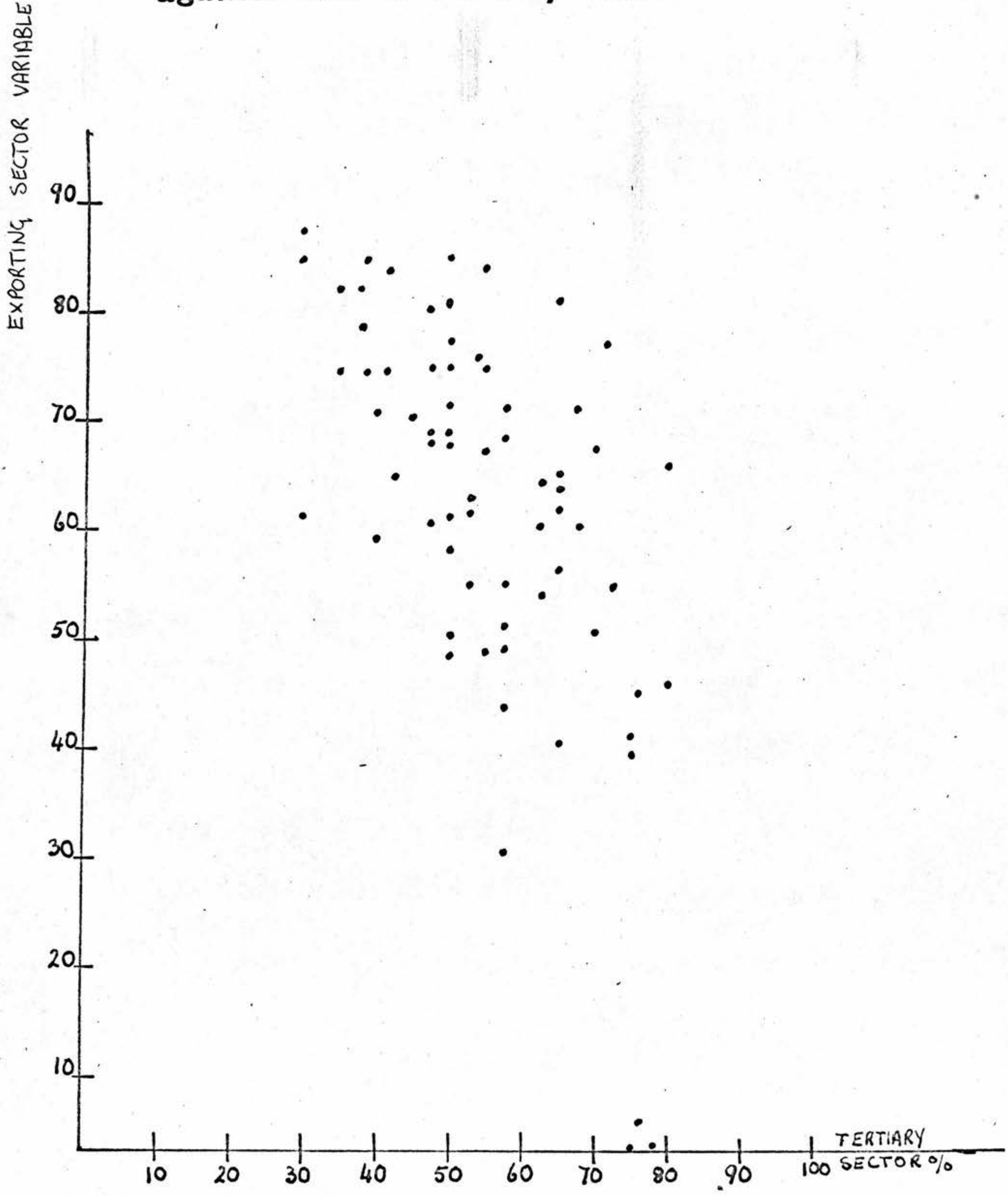
diagram C7

Index of Earnings in Secondary Sector against Tertiary Sector



diagramC8

Size of Exporting Sector against size of Tertiary Sector



APPENDIX D

THE SAMPLES

Table D 1

THE FULL SAMPLE

Aberdeen	Exeter	Scunthorpe
Barnsley	Gloucester	Sheffield
Barrow	Great Yarmouth	Southampton
Bath	Greenock	Southport
Bedford	Grimsby	Stoke
Blackburn	Harrogate	Sunderland
Blackpool	Hastings	Swansea
Bournemouth	Ipswich	Swindon
Brighton	Kingston-upon-Hull	Torquay
Bristol	Kirkcaldy	Wakefield
Burnley	Leicester	Warrington
Burton	Lincoln	West Hartlepool
Cambridge	Mansfield	Widnes
Cardiff	Merthyr Tydfil	Wigan
Carlisle	Newport	Worcester
Cheltenham	Northampton	Worthing
Chester	Norwich	York
Chesterfield	Nottingham	
Colchester	Nuneaton	
Coventry	Oxford	
Crewe	Peterborough	
Darlington	Plymouth	
Derby	Portsmouth	
Doncaster	Preston	
Dundee	Rugby	
Eastbourne		
Edinburgh		

Table D 2

THE 38, 55 and 69 SAMPLES

The 38 Sample

Aberdeen	Greenock	Plymouth
Barnsley	Grimsby	Portsmouth
Barrow	Ipswich	Preston
Bedford	Kirkcaldy	Southampton
Bristol	Lincoln	Southport
Burnley	Mansfield	Swansea
Burton	Merthyr Tydfil	Swindon
Cardiff	Newport	Wakefield
Chesterfield	Northampton	West Hartlepool
Crewe	Norwich	Wigan
Darlington	Nuneaton	Worcester
Doncaster	Oxford	York
Edinburgh	Peterborough	

The 55 Sample

The 38 sample plus:

Blackburn	Dundee	Rugby
Cambridge	Exeter	Stoke
Carlisle	Gloucester	Sunderland
Colchester	Kingston-upon-Hull	Warrington
Coventry	Leicester	Widnes
Derby	Nottingham	

The 69 Sample

The 55 sample plus:

Bath	Chester	Scunthorpe
Blackpool	Eastbourne	Sheffield
Bournemouth	Great Yarmouth	Torquay
Brighton	Harrogate	Worthing
Cheltenham	Hastings	

Table D 3

THE NORTH AND SOUTH SAMPLES

The North

Aberdeen	Dundee	Scunthorpe
Barnsley	Edinburgh	Sheffield
Barrow	Greenock	Southport
Blackburn	Grimsby	Stoke
Blackpool	Harrogate	Sunderland
Burnley	Kingston-upon-Hull	Wakefield
Carlisle	Kirkcaldy	Warrington
Chester	Lincoln	West Hartlepool
Chesterfield	Mansfield	Widnes
Crewe	Nottingham	Wigan
Darlington	Preston	York
Doncaster		

The South

Bath	Eastbourne	Oxford
Bedford	Exeter	Peterborough
Bournemouth	Gloucester	Portsmouth
Brighton	Great Yarmouth	Plymouth
Bristol	Hastings	Rugby
Burton	Ipswich	Southampton
Cambridge	Leicester	Swansea
Cardiff	Merthyr Tydfil	Swindon
Cheltenham	Newport	Torquay
Colchester	Northampton	Worthing
Coventry	Norwich	Worcester
Derby	Nuneaton	

Table D 4

THE COUNTIES

Cambridge	Lancashire	Suffolk
Cheshire	Norfolk	Sussex
Cumberland	Northampton	Warwickshire
Devon	Nottingham	Wiltshire
Durham	Oxford	Worcestershire
Glamorgan	Somerset	Yorkshire East
Gloucester	Stafford	Yorkshire North
Hampshire		

<u>APPENDIX E</u>

APPENDIX E

THE SPECIFICATION OF THE INDEPENDENT VARIABLES BASED UPON INDUSTRIAL STRUCTURE

This appendix is concerned with investigating the possibility that the negative coefficients of the export sector and index of earnings variables may have been imposed upon the analysis by the particular specification of those variables.

EXPORT SECTOR VARIABLE

In Chapter 5 this variable was defined in two stages. First, the percentages of the total labour force employed in those secondary industrial orders which were greater in size than the national average were aggregated. Secondly, this total was divided by the percentage of the labour force in the secondary sector.

If A represents the absolute employment in those industries in the secondary sector which are greater than the national average, then the absolute size of the secondary sector will be (A+B), where B is the absolute employment in those industrial orders smaller than the national average. If the employment in the tertiary sector is C, then the total labour force will be (A+B+C).

Using this notation, then the first step above results in the fraction $\frac{A}{A+B+C}$, whilst the second involves dividing this by $\frac{A+B}{A+B+C}$. Hence the export sector variable (E_t) is given by

$$\frac{A}{A+B+C} \Big/ \frac{A+B}{A+B+C} \quad \text{or} \quad \frac{A}{(A+B)}$$

Similarly, the size of the tertiary sector is $\frac{C}{A+B+C}$. Thus the hypothesis that the size of the tertiary sector is a function of the export sector variable can be re-written as

$$\frac{C}{A+B+C} = k + l\,\frac{A}{A+B} \qquad \text{(E.1)}$$

The question now to be asked is what conditions are necessary to produce a negative coefficient for L? This can be approached through investigating, first, the likely changes in A, B and C as the regression for equation (E.1) is performed and, secondly, of the conditions which must attach to these changes for a negative coefficient to be imposed upon L.

Clearly a considerable variety of factors must affect the sizes of A, B and C, but the most significant one is likely to be the town's population, since A, B and C are the absolute numbers employed in the relative sectors. Moreover it would be expected that a rise in population would cause A, B and C to rise.

Thus consider a rise in A consequent upon a rise in the town's population. For there to be a negative coefficient upon C, the factors $\frac{A}{A+B}$ and $\frac{C}{A+B+C}$ must move in opposite directions. The fraction $\frac{A}{A+B}$ will rise if the following condition is met, since A and B are both positive and greater than unity -

$$\%\text{ rise in A} > \%\text{ rise in B} \qquad \underline{\text{Condition One}}$$

Now the fraction $\frac{C}{A+B+C}$ will fall if

$$\%\text{ rise in C} < \%\text{ rise in (A+B)} \qquad \underline{\text{Condition Two}}$$

Exactly analogous conditions would be met (still considering a rise in A, B and C) for the fraction $\frac{A}{A+B}$ to fall and that of $\frac{C}{A+B+C}$ to rise.

Hence, for a negative coefficient to be imposed upon L, conditions one and two above are both necessary and sufficient. The question now is whether these conditions were in fact met in the particular sample adopted in the analysis.

When the relationship between the export sector variable and the population of the town was investigated, the correlation coefficient was found to be 0.032. This shows that the fraction $\frac{A}{A+B}$ does not alter with population, so that between towns of different population levels condition one cannot be true.

Similarly the relationship between the size of the tertiary sector and population produced a correlation coefficient of 0.062. Since $(A+B)$ does not alter with population, this would imply that the relationship between $(A+B)$ and C does not alter either. Hence condition two cannot be met in the particular sample adopted in this analysis.

Thus, subject to the initial assumption, viz. that an increase in population will cause an increase in A, B and C, these conditions are not met and there is no reason to suppose that the particular specification of the hypothesis forced a negative coefficient upon L.

Nevertheless it is important to note that such a sign could be produced under either of the two following conditions:

a) the rise in population was not sufficient to counteract other factors contributing to a fall in A, B or C. However, considering that A, B and C are absolute employment numbers this would be unlikely to occur in practice.

b) the population is constant so that other factors could have influenced A, B and C. This is but a limiting case of the above analysis based upon a change in population.

INDEX OF EARNINGS VARIABLE

Essentially the same conclusions can be arrived at for this variable through a similar type of analysis. The index of earnings variable was again computed in two stages. First the percentages of employment in each industrial order in the primary and secondary sectors [orders 1-17] were multiplied by the national wage earnings for that order and, secondly, the aggregate of these was divided by the percentage size of the primary and secondary sectors. If C is again the size of the tertiary sector and Z_i is the size of the individual primary or secondary orders, the first step results in the fraction $\dfrac{\sum_{i=1}^{17} z_i w_i}{Z + C}$ [where $\sum_{i=1}^{17} z_i = Z$ for ease of notation].

The second step requires this fraction to be divided by $\dfrac{Z}{Z + C}$. Hence, the index of earnings variable can be re-written as

$$T_t = \frac{\sum_{i=1}^{17} z_i w_i}{Z}$$

and the size of the tertiary sector is

$$S_t = \frac{C}{Z + C}$$

Thus the hypothesis to be tested is

$$\frac{C}{Z + C} = m + p \frac{\sum_{i=1}^{17} z_i w_i}{Z} \qquad (E.2)$$

Adopting the same assumptions for the movement of Z and C with population as in the previous section, a rise in z_i (and therefore Z) will cause T_t to rise, since w is greater than unity. Hence, for p to carry a negative coefficient, S_t must fall, which means that

% rise in C $<$ % rise in Z **Condition Three**

Yet, as already noted, the fraction $\frac{C}{Z+C}$ (i.e. $\frac{C}{A+B+C}$ in the previous section) does not vary with population. So, subject to the reservations made concerning the export sector variable, it can be stated that the particular specification of the index of earnings variable does not of necessity impose a negative coefficient upon the analysis.